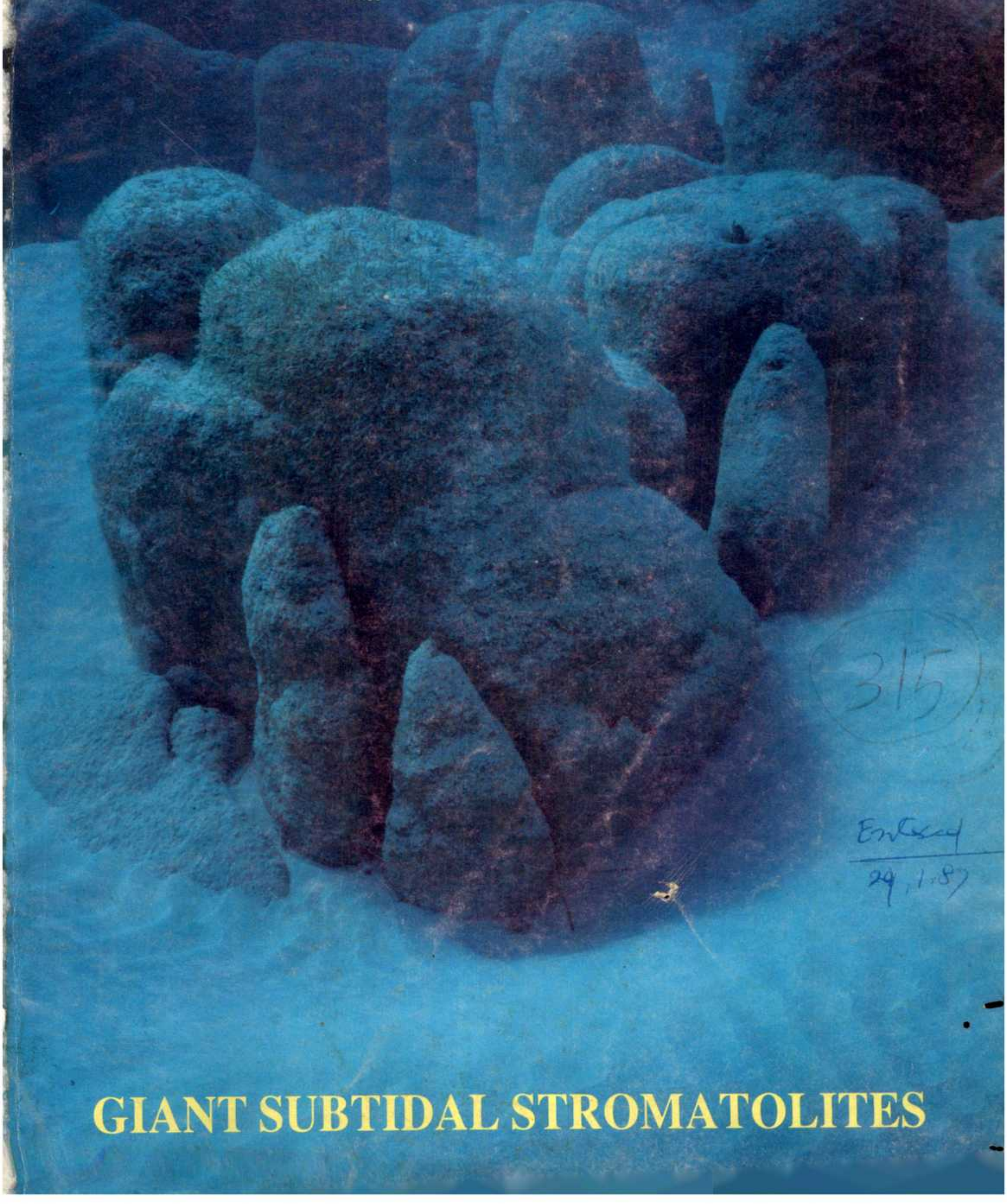
nature
INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
Volume 324 No.6092 6-12 November 1986 £1.90

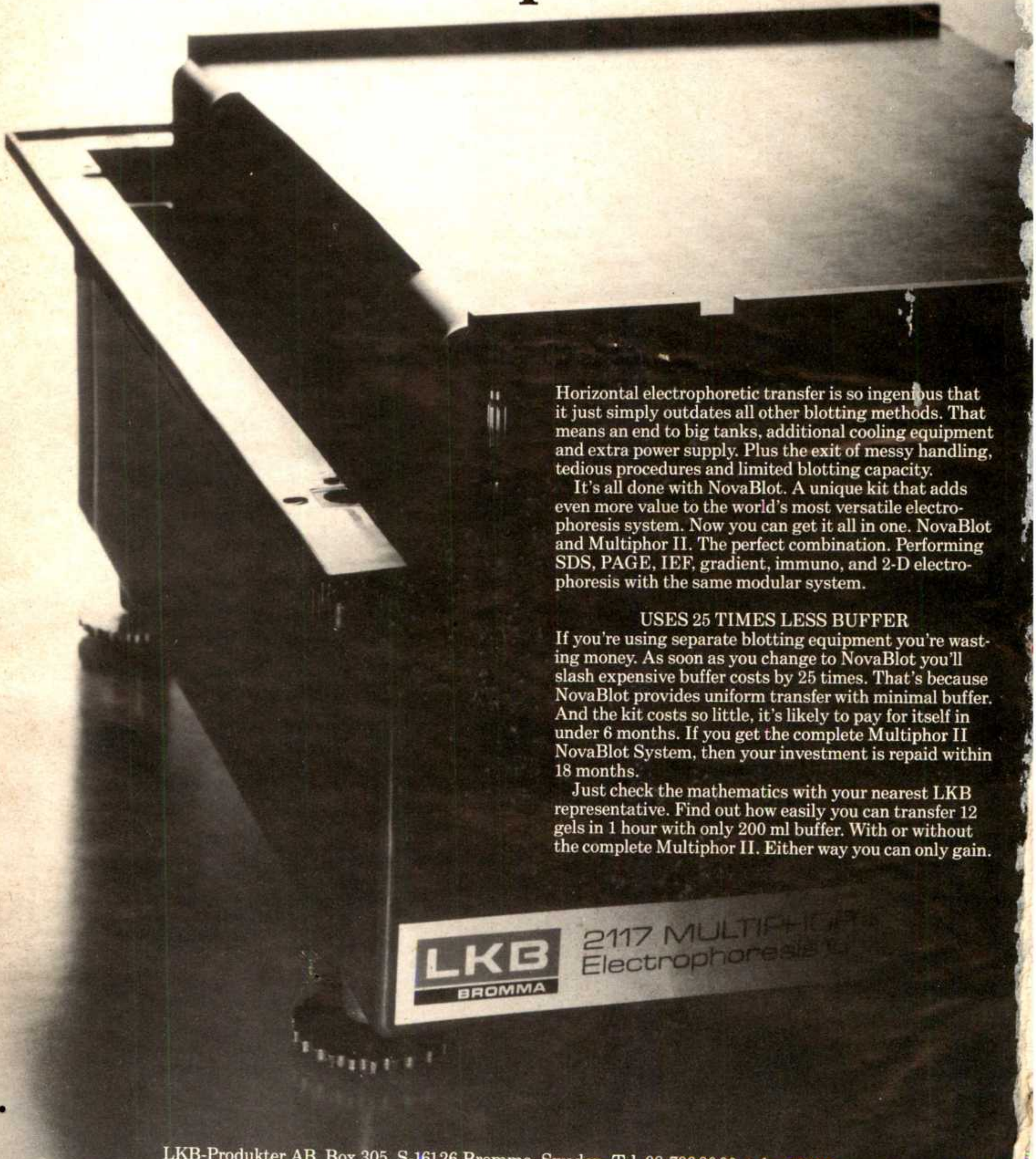


GIANT SUBTIDAL STROMATOLITES

315

Entered
29.1.87

Now Multiphor II gives you horizontal electrophoretic transfer



Horizontal electrophoretic transfer is so ingenious that it just simply outdates all other blotting methods. That means an end to big tanks, additional cooling equipment and extra power supply. Plus the exit of messy handling, tedious procedures and limited blotting capacity.

It's all done with NovaBlot. A unique kit that adds even more value to the world's most versatile electrophoresis system. Now you can get it all in one. NovaBlot and Multiphor II. The perfect combination. Performing SDS, PAGE, IEF, gradient, immuno, and 2-D electrophoresis with the same modular system.

USES 25 TIMES LESS BUFFER

If you're using separate blotting equipment you're wasting money. As soon as you change to NovaBlot you'll slash expensive buffer costs by 25 times. That's because NovaBlot provides uniform transfer with minimal buffer. And the kit costs so little, it's likely to pay for itself in under 6 months. If you get the complete Multiphor II NovaBlot System, then your investment is repaid within 18 months.

Just check the mathematics with your nearest LKB representative. Find out how easily you can transfer 12 gels in 1 hour with only 200 ml buffer. With or without the complete Multiphor II. Either way you can only gain.

LKB-Produkter AB, Box 305, S-16126 Bromma, Sweden, Tel. 08-799 80 00, telex 10492
Antwerp (03)218 93 35 • **Athens-Middle East** +30(1)894 73 96 • **Beijing** 89 06 21 • **Copenhagen** (01) 29 50 44 • **HongKong** (852) 5-814 84 21
London (01) 657 88 22 • **Lucerne** (041) 57 44 57 • **Madras** (044) 45 28 74 • **Moscow** (095) 255-6984 • **Munich** (089) 85 830
Paris (01) 64 46 36 36 • **Rome** (06) 39 90 33 • **Sao Paulo** (011) 578 41 38 • **Stockholm** (08) 764 75 50 • **Tokyo** (03) 293-5141
Turku (021) 678 111 • **Vienna** +43(222) 92 16 07 • **Washington** (301) 963 3200 • **Zoetermeer** (079) 31 92 01
Over 60 qualified representatives throughout the world.

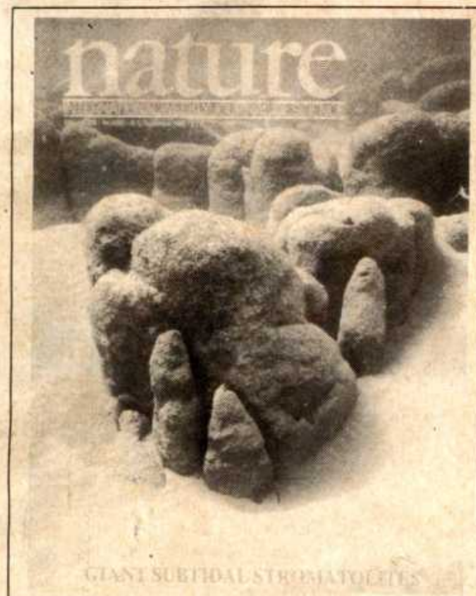
Reader Service No.40

Sporre Medical Inf

nature

315

NATURE VOL. 324 6 NOVEMBER 1986



Underwater view of giant stromatolites (up to 2.5 metres high forming at a depth of 7 metres in a tidal channel in the Bahamas. These are the largest modern examples yet discovered, and their occurrence in water of normal salinity sheds new light on the environmental significance of their ancient counterparts, which flourished in the Precambrian, more than 2,500 million years ago (see p.55).

OPINION

What must be done about AIDS 1

NEWS

AIDS 3
AIDS drug test
Test-ban seismology 4
British seek funds
Strategic Defense Initiative
Robots 5
Alaska glacier seals fate of fiord wildlife
US education
US oceanography 6
Soviet Union
Japanese physics 7
One-way crossing
What anniversary?
European research 8
British research
US cancer 9
French research

CORRESPONDENCE

A cosmic rod in pickle/ Indian collaboration/ Faraway quasars 10

NEWS AND VIEWS

Physicists about to hi-jack DNA? 11
Millimetre-sized quasicrystals grown from aluminium alloy
Kevin Knowles 12
Molecular biology: Manipulation just off target
R Michael Liskay 13
Elementary particles: Colour theory in a spin
Harry J Lipkin 14
Unravelling the control of B cells
Gerry G B Klaus 16
Questions of taste and smell
Gordon M Shepherd, Thomas V Getchell & Charlotte M Mistretta 17
New radiocarbon dating system
Roy Switsur 18
The environmentalist myth in archaeology
Jared M Diamond 19

SCIENTIFIC CORRESPONDENCE

Where does the message begin?
TE Creighton
Reply: S H Orkin 21
Occam and mankind's genetic bottleneck
J J D Greenwood 21
Chernobyl fallout in size-fractionated aerosol
D J Jost, H W Gäggeler, U Baltensperger, B Zinder & P Haller 22
Alimentary restrictions and ^{131}I in human thyroids
P Orlando, G Gallelli, F Perdelli, S DeFlora & R Malconteni 23
Radioactive fallout swipe samples from Chernobyl
S Sadasivan & U C Mishra 23

BOOK REVIEWS

Inventing the NIH: Federal Biomedical Research Policy, 1887–1937
by V A Harden
Alan Glynn 25
Human Brainwaves: The Psychological Significance of the Electroencephalogram
by J Empson
D Aled Jeffreys 26
Encyclopedia of Materials Science and Engineering
M B Bever, editor-in-chief
Charles Newey and colleagues 26
Introductory Biophysics
by M Cerdonio and R W Noble
Michael Holwill 27
The Age of Science
by D Knight
W H Brock 28
How to Deep-Freeze a Mammoth
by B Kurtén
Beverly Halstead 28

COMMENTARY

Cost-effectiveness of short-term tests for carcinogenicity
L B Lave & G S Omenn 29

ARTICLES

Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene
K R Thomas & M R Capecchi 34

LETTERS TO NATURE

Outbursts of Cygnus X3 observed at 1.3 and 3.3 mm wavelengths
J W M Baars, W J Altenhoff, H Hein & H Steppe 39
Io plasma torus and the source of jovian radiation (bKOM)
D Jones 40
An icy-glue model of cometary nuclei
T I Gombosi & H L F Houpis 43
Distribution of energetic particles near interplanetary shocks
A Hewish & S Bravo 44
Searching potential energy surfaces by simulated annealing
L T Wille 46
Large Al–Cu–Li single quasicrystals with triacontrahedral solidification morphology
B Dubost, J-M Lang, M Tanaka, P Sainfort & M Audier 48
Evidence for carbonate in the mantle
G W Berg 50
Lahars initiated by the 13 November 1985 eruption of Nevado del Ruiz, Colombia
D R Lowe, S N Williams, H Leigh, C B Connor, J B Gemmell & R E Stoiber 51
Sporadic shutdown of North Atlantic deep water production during the Glacial-Holocene transition?
W H Berger & E Vincent 53
Giant subtidal stromatolites forming in normal salinity waters
R F Dill, E A Shinn, A T Jones, K Kelly & R P Steinen 55
Sex-ratio manipulation in the common opossum
S N Austad & M E Sunquist 58

Contents continued overleaf

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Journals Ltd and includes the Annual Index (mailed in February). Annual subscription for USA and Canada US $250. USA and Canadian orders to: *Nature*, Subscription Dept, PO Box 1501, Neptune, NJ 07753, USA. Other orders to *Nature*, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Second class postage paid at New York, NY 10012 and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $1.00 a copy plus $0.10 a page is paid directly to CCC, 21 Congress Street, Salem, Massachusetts 01970, USA. Identification code for *Nature*: 0028-0836/86 $1.00 + $0.10. US Postmaster send address changes to: *Nature*, 65 Bleecker Street, New York, NY 10012. Published in Japan by Macmillan Shuppan K.K., Eikow Building, 1-10-9 Hongo, Bunkyo-ku, Tokyo, Japan. © 1986 Macmillan Journals Ltd.

CUSTOM SYNTHETIZED DNA AT EXTREMELY LOW PRICES

Prices (USD) per purified oligomer up to the length of 35:

	Basic fee	Per base
0.05 A260 units (~ 1.5 μg)	85	5.50
0.2 A260 units (~ 6 μg)	125	7.50
1 A260 unit (~30 μg)	170	10.00

SPECIAL OFFER: FREE OLIGONUCLEOTIDES

FOR EVERY FIVE OLIGOMERS YOU BUY
YOU GET ONE ADDITIONAL OLIGOMER FOR FREE

WHY BUY MORE THAN YOU ACTUALLY NEED?

1 μg OF OLIGONUCLEOTIDE IS SUFFICIENT FOR:

- 10 oligonucleotide-directed mutagenisis experiments.
- 100 hybridization and gene synthesis experiments.
- 1000 dideoxy sequencing experiments.
- More than 1000 cloning experiments.

RP442

EVALUATE YOUR ACTUAL NEEDS BEFORE ORDERING!

FOR PRICE LIST AND PRODUCT INFORMATION, CONTACT:

SYN-TEK AB
Box 1010
S-901 20 Umeå, SWEDEN
Phone +46-90-13 24 10
Telefax +46-90-12 32 05
Telex 54920 syntek s

SYN-TEK CORP.
P.O. Box 50068
Palo Alto, CALIFORNIA 94303
Phone 415-494-8690
Telefax 415-494-0762

SYN-TEK AB
INNOVATORS IN BIOLOGICAL SCIENCES
Reader Service No. 20

CONTENTS

NATURE BINDERS

Your copies of Nature are invaluable for future reference so why not give them the treatment they deserve and file them neatly in our special Nature binders?

Ordering address and prices for one or a set of four are listed below.

Annual Subscription Prices

UK		£104
Rest of Europe	Airspeed	£125
USA & Canada		US$250
Australia, NZ & S. Africa	Airspeed	£150
	Airmail	£190
India	Airspeed	£120
	Airmail	£190
Japan	Airspeed	Y90000
Rest of World	Surface	£120
	Airmail	£180

Orders (with remittance) to:

USA & Canada
Nature
Subscription Dept
PO Box 1501
Neptune
NJ 07753
USA

UK & Rest of World
Nature
Circulation Dept
Brunel Road
Basingstoke
Hants RG21 2XS, UK
Tel: 0256 29242

(The addresses of Nature's editorial offices are shown facing the first editorial page)

Japanese subscription enquiries to:
Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan
Tel: (03) 292 3755

Personal subscription rates: These are available in some countries to subscribers paying by personal cheque or credit card. Details from:

USA & Canada
Nature
65 Bleecker Street
New York
NY 10012, USA
Tel: (212) 477-9600

UK & Europe
Felicity Parker
Nature
4 Little Essex Street
London WC2R 3LF, UK
Tel: 01-836 6633

Back issues: UK, £2.50; USA & Canada, US$6.00 (surface), US$9.00 (air); Rest of World, £3.00 (surface), £4.00 (air)

Binders
Single binders: UK, £5.25; USA & Canada $11.50; Rest of World, £7.75. Set of four: UK, £15.00; USA & Canada $32.00; Rest of World, £22.00

Annual indexes (1971-1985)
UK, £5.00 each; Rest of World, $10.00

Nature First Issue Facsimile
UK, £2.00; Rest of World (surface), $3.00; (air) $4.00

Nature in microform
Write to University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA

Wormy mice in a hybrid zone
R D Sage, D Heyneman, Kee-Chong Lim & A C Wilson 60

A single micromanipulated stem cell gives rise to multiple T-cell receptor gene rearrangements in the thymus *in vitro*
G T Williams, R Kingston, M J Owen, E J Jenkinson & J J T Owen 63

Extensive proliferation of mature connective-tissue type mast cells *in vitro*
T Nakahata, T Kobayashi, A Ishiguro, K Tsuji, K Naganuma, O Ando, Y Yagi, K Tadokoro & T Akabane 65

A chimaeric receptor allows insulin to stimulate tyrosine kinase activity of epidermal growth factor receptor
H Riedel, T J Dull, J Schlessinger & A Ullrich 68

Cloning of complementary DNA encoding T-cell replacing factor and identity with B-cell growth factor II
T Kinashi, N Harada, E Severinson, T Tanabe, P Sideras, M Kinishi, C Azuma, A Tominaga, S Bergstedt-Lindqvist, M Takahashi, F Matsuda, Y Yaoita, K Takatsu & T Honjo 70

Complementary DNA for a novel human interleukin (BSF-2) that induces B lymphocytes to produce immunoglobulin
T Hirano, K Yasukawa, H Harada, T Taga, Y Watanabe, T Matsuda, S Kashiwamura, K Nakajima, K Koyama, A Iwamatsu, S Tsunasawa, F Sakiyama, H Matsui, Y Takahara, T Taniguchi & T Kishimoto 73

Intracellularly injected tetanus toxin inhibits exocytosis in bovine adrenal chromaffin cells
R Penner, E Neher & F Dreyer 76

The essential light chains constitute part of the active site of smooth muscle myosin
Y Okamoto, T Sekine, J Grammer & R G Yount 78

Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases
P K Hwang & R J Fletterick 80

α-Lactalbumin possesses a novel calcium binding loop
D I Stuart, K R Acharya, N P C Walker, S G Smith, M Lewis & D C Phillips 84

Bent DNA at a yeast autonomously replicating sequence
M Snyder, A R Buchman & R W Davis 87

Flavonoid activation of nodulation genes in *Rhizobium* reversed by other compounds present in plants
J L Firmin, K E Wilson, L Rossen & A W B Johnston 90

Trans-sexually grafted antennae alter pheromone-directed behaviour in a moth
A M Schneiderman, J G Hildebrand, M M Brennan & J H Tumlinson **Vol. 323, 801**
(This letter was omitted from the contents page of the 30 October issue.)

NEW ON THE MARKET

Notes for the laboratory 93

EMPLOYMENT

The graduate labour market
R Pearson 94

MISCELLANY

100 years ago 19

NATURE CLASSIFIED

Professional appointments — Research posts — Studentships — Fellowships — Conferences — Courses — Seminars — Symposia: Back Pages

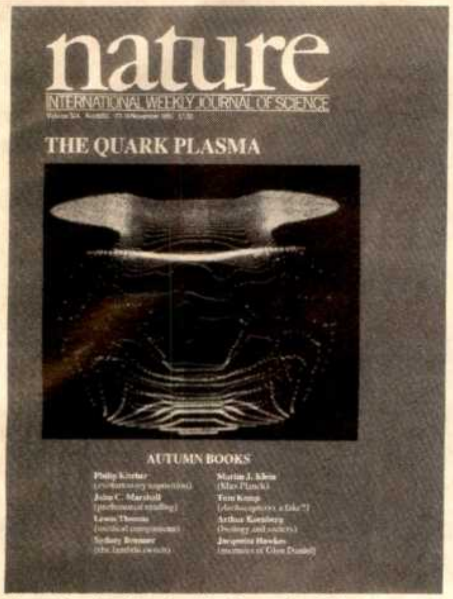
Next week in *Nature*:

- Colonizing space
- Companion for millisecond pulsar
- New polymer fuel
- Search for the quark plasma
- New hominids from Kenya
- Plant haemoglobins
- Fragile X
- Autumn Books
- Mass extinction of terrestrial plants
- Divergent genetic codes
- Dynamic microtubules

ALL THE CHAINS YOU NEED.

nature

NATURE VOL. 324 13 NOVEMBER 1986



Very shortly after the Big Bang, the Universe was a plasma of three quarks. Quantum chromodynamics has provided first predictions for such quark matter, and an experimental quark plasma search is starting this autumn using energetic nuclear collisions to produce bubbles of 'primordial matter'. The cover shows a computer simulation of such a collision between oxygen and uranium nuclei. See page 116.

OPINION

What price independence? 95
Research in trouble 96

NEWS

British penury relaxed 97
Meningitis
US mid-term elections
Strategic defence
British SDI cash
Canadian budgets cut 99
Accord renewed
New chairman for vice-chancellors
Hubble telescope taking shape 100
Coral reefs
Yugoslavia 101
X-ray astronomy
Ups and downs of French budget
Semiconductor rivalry 102
Ozone depletion
Deaths foretold
US agricultural research 103
Grassroots unrest in plant biology
FRG-Soviet agreement

CORRESPONDENCE

Rescuing British science/etc. 104

NEWS AND VIEWS

How to tell when the sea rises? 105
Dynamic microtuble dynamics
Robley C Williams Jr, Michael Caplow & J Richard McIntosh 106
Rapid evolution of urban birds
Jared M Diamond 107
The supermodulus syndrome
Robert W Cahn 108
Molecular evolution: Deciphering divergent codes
L A Grivell 109
Moon's birth shocked Earth?
Alan P Boss 110
Fusion energy: Highest temperatures yet
R S Pease 111
Catastrophic vegetation changes
Margaret E Collinson 112

SCIENTIFIC CORRESPONDENCE

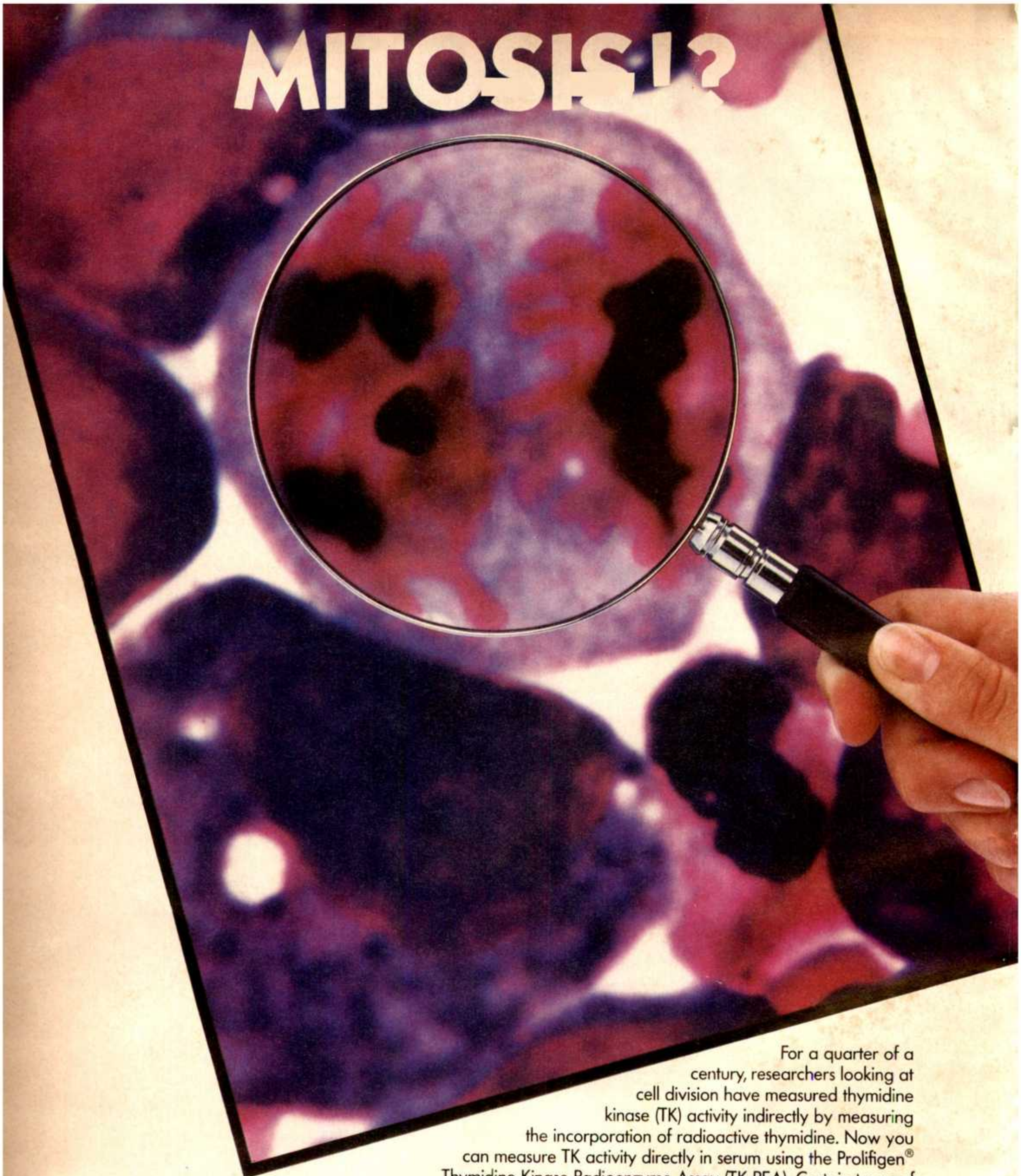
Dating and correlating the French Mousterian
N Ashton & J Cook
Reply: P Mellars 113
What can AIDS virus codon usage tell us?
P M Sharp 114

AUTUMN BOOKS

Study in indifference
John Manger on the trans-Atlantic contrasts in the reception accorded to a biography of Albert Einstein 169
The Blind Watchmaker
by R Dawkins
Philip Kitcher 170
Storm Over Biology: Essays on Science, Sentiment, and Public Policy
by B D Davis
Arthur Kornberg 172
The Problem of Altruism: Freudian–Darwinian Solutions
by C R Badcock
John R Durant 172
The Monkey Gland Affair
by D Hamilton
Walter Gratzer 173
The Oxford Companion to Medicine
J Walton, P B Beeson and R Bodley Scott, eds
Lewis Thomas 174
Sewall Wright and Evolutionary Biology *by W B Provine* and
Evolution: Selected Papers by Sewall Wright *W B Provine, ed.*
Joe Felsenstein 175
Wonders of Creation: Natural History Drawings in the British Library *by R Desmond* and
A Redouté Treasury
text by P and F Mallary
Sandra Raphael 176
The History of Scurvy and Vitamin C
by K J Carpenter
John Rivers 177
Molecular Biology of Development
John C Gerhart 178
Gilbert White: A Biography of the Author of The Natural History of Selborne *by R Mabey*
Robert M May 179
A Genetic Switch: Gene Control and Phage λ *by M Ptashne*
Sydney Brenner 179
China: Land of Discovery and Invention *by R K G Temple*
Christopher Cullen 180
False Prophets: Fraud and Error in Science and Medicine
by A Kohn
C R B Joyce 181
Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer
by H L Dreyfus and S E Dreyfus
Steve Woolgar 182
Nature's Gambit: Child Prodigies and the Development of Human Potential
by D H Feldman with L T Goldsmith
Peter Bryant 183
The Tenth Muse: The Pursuit of Earth Science *by R B Parker*
Derek Ager 184
Islands *by H W Menard*
Peter J Smith 184
Archaeopteryx, The Primordial Bird: A Case of Fossil Forgery
by F Hoyle and C Wickramasinghe
Tom Kemp 185
Mammal Evolution: An Illustrated Guide
by R J G Savage and M R Long
Malcolm C McKenna 186
Controversy in Victorian Geology: The Cambrian–Silurian Dispute
by J A Secord
A Hallam 186
The Harvard Guide to Influential Books *C M Devine, C M Dissel and K D Parrish, eds*
John C Marshall 187
Some Small Harvest: The Memoirs of Glyn Daniel
Jacquetta Hawkes 188
The Palaeolithic Settlement of Europe
by C Gamble
Randall White 189
The Dilemmas of an Upright Man: Max Planck as Spokesman for German Science *by J L Heilbron*
Martin J Klein 190
Was Einstein Right? Putting General Relativity to the Test *by C M Will*
Joseph H Taylor 191

Contents continued overleaf

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Journals Ltd and includes the Annual Index (mailed in February). Annual subscription for USA and Canada US $250. USA and Canadian orders to: *Nature*, Subscription Dept, PO Box 1501, Neptune, NJ 07753, USA. Other orders to *Nature*, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Second class postage paid at New York, NY 10012 and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $1.00 a copy plus $0.10 a page is paid directly to CCC, 21 Congress Street, Salem, Massachusetts 01970, USA. Identification code for *Nature*: 0028-0836/86 $1.00 + $0.10. US Postmaster send address changes to: *Nature*, 65 Bleecker Street, New York, NY 10012. Published in Japan by Macmillan Shuppan K.K., Eikow Building, 1-10-9 Hongo, Bunkyo-ku, Tokyo, Japan. © 1986 Macmillan Journals Ltd



For a quarter of a century, researchers looking at cell division have measured thymidine kinase (TK) activity indirectly by measuring the incorporation of radioactive thymidine. Now you can measure TK activity directly in serum using the Prolifigen® Thymidine Kinase Radioenzyme Assay (TK-REA). Certain types of malignant disease are characterised by a particularly rapid cell proliferation; increased proliferation means an increased level of TK. Serum TK measurements have a proven clinical use for prognosis and follow-up in such diseases, particularly lymphomas, leukaemias and small-cell lung cancer. Contact Sangtec today and measure TK the easy way.

TK-REA is the way to measure Thymidine Kinase.

Sangtec MEDICAL

In U.S.A. for research purposes only

Box 20045, S-161 20 Bromma, Sweden Tel. Int + 468 988330

Reader Service No.47

CONTENTS

In this week's issue...

How you can make a Christmas Present of a year's subscription to Nature and receive a free gift!

Look inside for further information or ring Za Dale on (01) 836 6633

Annual Subscription Prices

UK		£104
Rest of Europe	Airspeed	£125
USA & Canada		US$250
Australia, NZ & S. Africa	Airspeed	£150
	Airmail	£190
India	Airspeed	£120
	Airmail	£190
Japan	Airspeed	Y90000
Rest of World	Surface	£120
	Airmail	£180

Orders (with remittance) to:

USA & Canada
Nature
Subscription Dept
PO Box 1501
Neptune
NJ 07753
USA

UK & Rest of World
Nature
Circulation Dept
Brunel Road
Basingstoke
Hants RG21 2XS, UK
Tel: 0256 29242

(The addresses of Nature's editorial offices are shown facing the first editorial page)

Japanese subscription enquiries to:
Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan
Tel: (03) 292 3755

Personal subscription rates: These are available in some countries to subscribers paying by personal cheque or credit card. Details from:

USA & Canada
Nature
65 Bleecker Street
New York
NY 10012, USA
Tel: (212) 477-9600

UK & Europe
Felicity Parker
Nature
4 Little Essex Street
London WC2R 3LF, UK
Tel: 01-836 6633

Back issues: UK, £2.50; USA & Canada, US$6.00 (surface), US$9.00 (air); Rest of World, £3.00 (surface), £4.00 (air)

Binders
Single binders: UK, £5.25; USA & Canada $11.50; Rest of World, £7.75. Set of four: UK, £15.00; USA & Canada $32.00; Rest of World, £22.00

Annual indexes (1971-1985)
UK, £5.00 each; Rest of World, $10.00

Nature First Issue Facsimile
UK, £2.00; Rest of World (surface), $3.00; (air) $4.00

Nature in microform
Write to University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA

Frontiers of Learning: The Ph.D. Octopus *by P Hesseling*
Nicholas A Rupke 192

The Shaky Game: Einstein, Realism, and the Quantum Theory *by Arthur Fine*
Philip Pearle 193

Shelter Island II: Proceedings of the 1983 Shelter Island Conference on Quantum Field Theory and the Fundamental Problems of Physics
R Jackiw et al., *eds*
Ya B Zeldovich 194

COMMENTARY

Colonization of space: destiny or folly
V.R. Eshleman 115

REVIEW ARTICLE

The quark plasma
H Satz 116

ARTICLES

Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*
H G Frohnhöfer & C Nüsslein-Volhard 120

LETTERS TO NATURE

A parallax effect due to gravitational micro-lensing
B Grieger, R Kayser & S Refsdal 126

Companion star of a binary millisecond pulsar
G A Wright & E D Loh 127

An explanation for fast downflows on the Sun
A N McClymont & I J D Craig 128

Poly(styrene peroxide): an autocombustible polymer fuel
K Kishore & T Mukundan 130

Hydrothermal methane plumes in the Mariana back-arc spreading centre
Y Horibe, K-R Kim & H Craig 131

Structure of the Kerguelen Plateau province from Seasat altimetry and seismic reflection data
M F Coffin, H L Davis & W F Haxby 134

Oxygen isotopes and sea level
J Chappell & N J Shackleton 137

Predicting aqueous aluminium concentrations in natural waters
C S Cronan, W J Walker & P R Bloom 140

A new Miocene hominoid from Kenya
R E Leakey & M G Leakey 143

A second new Miocene hominoid from Kenya
R E Leakey & M G Leakey 146

Vegetation, climatic and floral changes at the Cretaceous-Tertiary boundary
J A Wolfe & G R Upchurch Jr 148

Growth and longevity of krill during the first decade of pelagic whaling
A A Rosenberg, J R Beddington & M Basson 152

Ocular dominance shift in kitten visual cortex caused by imbalance in retinal electrical activity
B Chapman, M D Jacobson, H O Reiter & M P Stryker 154

Intracellular calcium ions decrease the affinity of the GABA receptor
M Inoue, Y Oomura, T Yakushiji & N Akaike 156

The t(8; 14) chromosomal translocation occurring in B-cell malignancies results from mistakes in V-D-J joining
F G Haluska, S Finver, Y Tsujimoto & C M Croce 158

Implications of fragile X expression in normal males for the nature of the mutation
D H Ledbetter, S A Ledbetter & R L Nussbaum 161

Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes
R K Saiki, T L Bugawan, G T Horn, K B Mullis & H A Erlich 163

Common evolutionary origin of legume and non-legume plant haemoglobins
J Landsmann, E S Dennis, T J V Higgins, C A Appleby, A A Kortt & W J Peacock 166

NATURE CLASSIFIED

Professional appointments — Research posts — Studentships — Fellowships — Conferences — Courses — Seminars — Symposia: **Back Pages**

Next week in *Nature*:

- Reviving a prize idea
- Smoke in 'nuclear winter'
- Burnt-out galaxies
- Neptune's shape
- Perceiving motion
- Glycine in neurotransmission
- Interleukins: alpha versus beta
- Another T-cell population
- Oncogene c-*myc* and liver cancer
- Pre-mRNA splicing

What's the difference between here,
and here?
ZURICH

Nothing…

In short, you'll find the same excellent standards of service and efficiency down on the ground at Zurich and Geneva airports as you're used to on a Swissair flight at 39,000 feet.

There are always times when you can't fly direct to your long-haul destination and a transit stop can be a daunting prospect.

At Zurich and Geneva you'll experience first rate amenities and an organisation that ensures that everything runs like clockwork. So that your short time on the ground becomes a pleasure, not a pain.

Connecting flights will take you to over 100 destinations worldwide, on which you can experience the level of service that's won us Business Traveller Magazine's Best Airline award every year since it started in 1980.

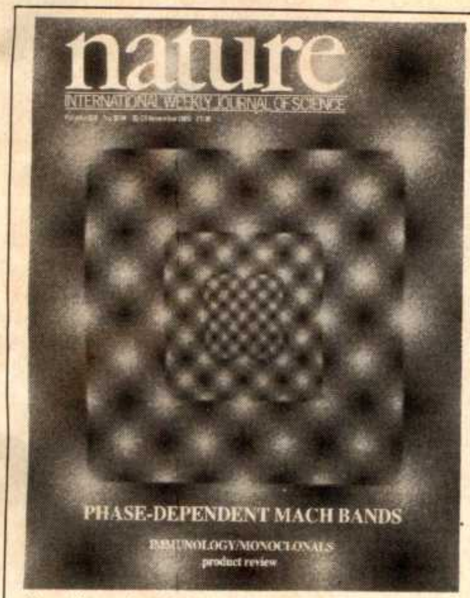
Fly Swissair. We make all the right connections.

swissair

Reader Service No.58

nature

NATURE VOL. 324 20 NOVEMBER 1986


nature
INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
PHASE-DEPENDENT MACH BANDS
IMMUNOLOGY/MONOCLONALS
product review

Two-dimensional Mach bands at four different scales. The horizontal and vertical cross-sections of the luminance distribution are triangular, but give rise to illusory star patterns. A theory of feature detection based on phase congruence (p.250) explains the appearance of the stars and why the pattern outline pinches in at smaller scales.

OPINION

Other people's nuclear weapons 195
Carnegie's misdiagnosis 196
Fire without smoke

NEWS

Research publishing 197
British space
Soviet accusation 198
Australian universities
Nuclear energy 199
AIDS
Storm flow simulation by computer
Technology 200
Student squeeze in West Germany
Chinese – US collaboration
Rhine pollution 201
Genetic manipulation 202
India
Japan 203
SERC annual report
Nuclear safety

CORRESPONDENCE

Creationism/Sexism/Letter from Moscow 204

NEWS AND VIEWS

Mathematics for natural hierarchy 205
National Parks and ecological realities
David S Wilcove & Robert M May 206
Superstrings: In search of a smoking gun
Roberto Peccei 207
Data bank for carbohydrate structures
Sarah Hargreaves 208
Arcs around Neptune
Carl D Murray 209
Squeezed states of light
Dan Walls 210
Plant biochemistry: Discovery of the most abundant inhibitor in the world?
Christine Foyer 211
Community ecology: Patterns in three dimensions
Paul H Harvey & John H Lawton 212
How fast does Halley spin?
Philip Campbell 213
Gene control mechanisms
Ian W Dawes 214

SCIENTIFIC CORRESPONDENCE

Fingers and helices
R S Brown & P Argos 215
Data in graphs and tables
R Ferreira 215
Silver from Chernobyl
H van Dam 216
Chernobyl fallout in Debrecen, Hungary
E Csongor, A Z Kiss, B M Nyakó & E Somorjai 216

BOOK REVIEWS

Intellectual Mastery of Nature: Theoretical Physics from Ohm to Einstein *by C Jungnickel and R McCormmach*
Peter Harman 217
Principles of Charged Particle Acceleration *by S Humphries Jr*
P W Hawkes 218
Accuracy in Molecular Processes: Its Control and Relevance to Living Systems *T B L Kirkwood, R F Rosenberger and D J Galas, eds*
Leslie E Orgel 219
Physics, Technology and Use of Photovoltaics *by R J Van Overstraeten and R P Mertens*
A K Turner 219
Introduction to the Cellular and Molecular Biology of Cancer *L M Franks and N Teich, eds*
D G Harnden 220
Science in Egyptology *A R David, ed*
D M Dixon 220

COMMENTARY

Glittering prizes for research support
D F Horrobin 221

REVIEW ARTICLE

Uncertainties in the smoke source term for 'nuclear winter' studies
J E Penner 222

ARTICLES

Occultation determination of Neptune's oblateness and stratospheric methane mixing ratio
E Lellouch, W B Hubbard, B Sicardy, F Vilas & P Bouchet 227

LETTERS TO NATURE

Have burnt-out galaxies and galaxy clusters been detected?
J Silk 231
A radio nebula associated with Circinus X-1
R F Haynes, M M Komesaroff, A G Little, D L Jauncey, J L Caswell, D K Milne, M J Kesteven, K J Wellington & R A Preston 233
The historical record for Sirius: evidence for a white-dwarf thermonuclear runaway?
F C Bruhweiler, Y Kondo & E M Sion 235
Ice core record of the $^{13}C/^{12}C$ ratio of atmospheric CO_2 in the past two centuries
H Friedli, H Lötscher, H Oeschger, U Siegenthaler & B Stauffer 237
Seasonal variation of methylarsenic compounds in airborne particulate matter
H Mukai, Y Ambe, T Muku, K Takeshita & T Fukuma 239
Reversed magnetization in pyroclastics from the 1985 eruption of Nevado del Ruiz, Colombia
F Heller, J C Carracedo & V Soler 241
Geochemical correlation between southern African kimberlites and South Atlantic hotspots
A P le Roex 243
Archaean flow-top alteration zones formed initially in a low-temperature sulphate-rich environment
D R Lowe & G R Byerly 245

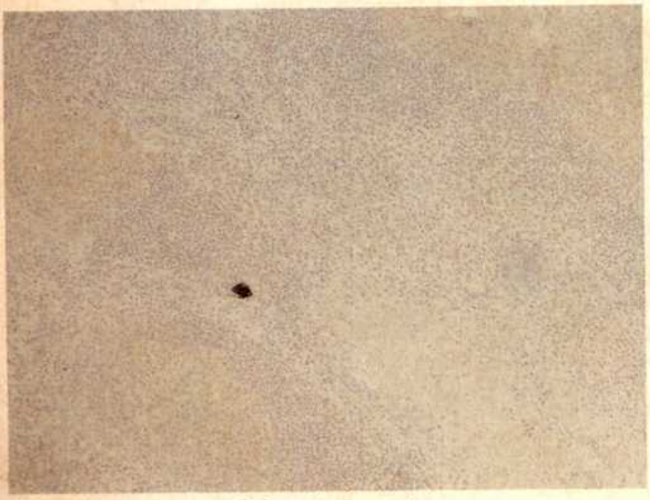
Contents continued overleaf

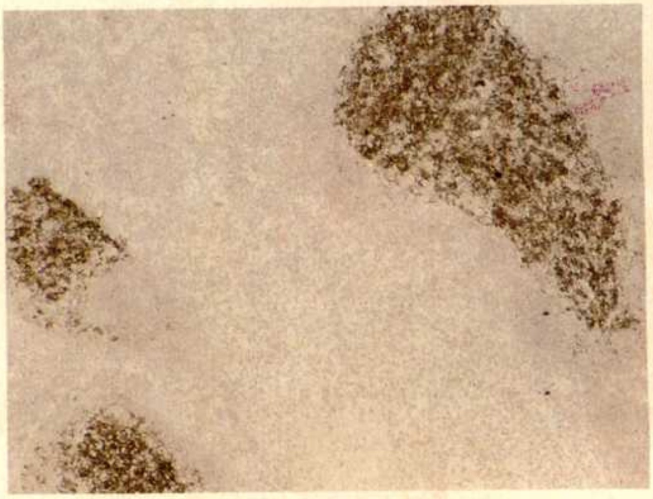
Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Journals Ltd and includes the Annual Index (mailed in February). Annual subscription for USA and Canada US $250. USA and Canadian orders to: *Nature*, Subscription Dept, PO Box 1501, Neptune, NJ 07753, USA. Other orders to *Nature*, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Second class postage paid at New York, NY 10012 and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $1.00 a copy plus $0.10 a page is paid directly to CCC, 21 Congress Street, Salem, Massachusetts 01970, USA. Identification code for *Nature*: 0028-0836/86 $1.00 + $0.10. US Postmaster send address changes to: *Nature*, 65 Bleecker Street, New York, NY 10012. Published in Japan by Macmillan Shuppan K.K., Eikow Building, 1-10-9 Hongo, Bunkyo-ku, Tokyo, Japan. © 1986 Macmillan Journals Ltd.

Enhance your DAB

with our DAB silver enhancement kit

A method for amplifying small amounts of 3,3′-diaminobenzidine precipitated on cells or tissues by immunoperoxidase.*


Before enhancement


After enhancement

- Immunoperoxidase staining of rare antigens can now be visualized.
- Concentrations of scarce or expensive monoclonal antibodies can be reduced without loss of sensitivity.
- Incubation times can be reduced.
- Improved contrast of immunoperoxidase-DAB photomicrographs.

*Patent applied for

Amersham International plc
Amersham Place Little Chalfont
Buckinghamshire England HP7 9NA
telephone Little Chalfont (02404) 4444
telex 838818 ACTIVL G

Amersham

UK Sales Aylesbury (0296) 35222 **Amersham Australia PTY Ltd** Sydney (2) 211 4399/4489 **Amersham Belgium SA/NV** Brussels (2) 538 9194
Amersham Buchler GmbH & Co KG Braunschweig, West Germany (05307) 8080 **Amersham Canada Ltd** Ontario (416) 842 2720
Amersham Corporation Arlington Heights USA (312) 364 7100 **Amersham Denmark ApS** Birkerød (02) 82 0222 **Amersham France SA** Paris (6) 928 8300
Amersham Japan Tokyo (3) 816 6161 **Amersham Nederland BV** Houten 03403 76660 **Amersham Sweden AB** Solna (8) 734 0800

Reader Service No.5

CONTENTS

In this week's issue...

How you can make a Christmas Present of a year's subscription to Nature and receive a free gift!

Look inside for further information or ring Za Dale on (01) 836 6633

Annual Subscription Prices

UK		£104
Rest of Europe	Airspeed	£125
USA & Canada		US$250
Australia, NZ & S. Africa	Airspeed	£150
	Airmail	£190
India	Airspeed	£120
	Airmail	£190
Japan	Airspeed	Y90000
Rest of World	Surface	£120
	Airmail	£180

Orders (with remittance) to:

USA & Canada
Nature
Subscription Dept
PO Box 1501
Neptune
NJ 07753
USA

UK & Rest of World
Nature
Circulation Dept
Brunel Road
Basingstoke
Hants RG21 2XS, UK
Tel: 0256 29242

(The addresses of Nature's editorial offices are shown facing the first editorial page)

Japanese subscription enquiries to:
Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan
Tel: (03) 292 3755

Personal subscription rates: These are available in some countries to subscribers paying by personal cheque or credit card. Details from:

USA & Canada
Nature
65 Bleecker Street
New York
NY 10012, USA
Tel: (212) 477-9600

UK & Europe
Felicity Parker
Nature
4 Little Essex Street
London WC2R 3LF, UK
Tel: 01-836 6633

Back issues: UK, £2.50; USA & Canada, US$6.00 (surface), US$9.00 (air); Rest of World, £3.00 (surface), £4.00 (air)

Binders
Single binders: UK, £5.25; USA & Canada $11.50; Rest of World, £7.75. Set of four: UK, £15.00; USA & Canada $32.00; Rest of World, £22.00

Annual indexes (1971-1985)
UK, £5.00 each; Rest of World, $10.00

Nature First Issue Facsimile
UK, £2.00; Rest of World (surface), $3.00; (air) $4.00

Nature in microform
Write to University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA

Body size, ecological dominance and Cope's rule
J H Brown & B A Maurer 248

Mach bands are phase dependent
M C Morrone, J Ross, D C Burr & R Owens 250

Hysteresis in the perception of motion direction as evidence for neural cooperativity
D Williams, G Phillips & R Sekuler 253

Inhibitory neurones of a motor pattern generator in *Xenopus* revealed by antibodies to glycine
N Dale, O P Ottersen, A Roberts & J Storm-Mathisen 255

T-cell epitope of the autoantigen myelin basic protein that induces encephalomyelitis
S S Zamvil, D J Mitchell, A C Moore, K Kitamura, L Steinman & J B Rothbard 258

Interaction of peptide antigens and class II major histocompatibility complex antigens
J-G Guillet, Ming-Zong Lai, T J Briner, J A Smith & M L Gefter 260

Identification of a common class of high affinity receptors for both types of porcine interleukin-1 on connective tissue cells
T A Bird & J Saklatvala 263

The cell surface receptors for interleukin-1α and interleukin-1β are identical
S K Dower, S R Kronheim, T P Hopp, M Cantrell, M Deeley, S Gillis, C S Henney & D L Urdal 266

Presence of Ti (WT31) negative T lymphocytes in normal blood and thymus
L L Lanier & A Weiss 268

Structure of pre-pro-von Willebrand factor and its expression in heterologous cells
D T Bonthron, R I Handin, R J Kaufman, L C Wasley, E C Orr, L M Mitsock, B Ewenstein, J Loscalzo, D Ginsburg & S H Orkin 270

A nocturnal inhibitor of carboxylation in leaves
S Gutteridge, M A J Parry, S Burton, A J Keys, A Mudd, J Feeney, J C Servaites & J Pierce 274

Rearrangement and enhanced expression of c-*myc* in hepatocellular carcinoma of hepatitis virus infected woodchucks
T Möröy, A Marchio, J Etiemble, C Trépo, P Tiollais & M-A Buendia 276

The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA
L P Eperon, J P Estibeiro & I C Eperon 280

Errata . . . page 282

A molecular link between the bats of New Zealand and South America
E D Pierson, V M Sarich, J M Lowenstein, M J Daniel & W E Rainey

Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells
S E Stachel, B Timmerman & P Zambryski

PRODUCT REVIEW

Immunoglobulin-derived drugs
Gary S Hahn 283

Immunotech ideas 284

NATURE CLASSIFIED

Professional appointments — Research posts — Studentships — Fellowships — Conferences — Courses — Seminars — Symposia: Back Pages

Next week in *Nature*:

- Science in Iberia
- Superfluid helium
- Origin of short-period comets
- New theropod dinosaur
- β-lactoglobulin structure

GUIDE TO AUTHORS

Authors should be aware of the diversity of *Nature*'s readership and should strive to be as widely understood as possible.

Review articles should be accessible to the whole readership. Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable).

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding. The text should not exceed 3,000 words and six displayed items (figures plus tables). The article should include an italic heading of about 50 words.

Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items. The first paragraph (not exceeding 150 words) should say what the letter is about, why the study it reports was undertaken and what the conclusions are.

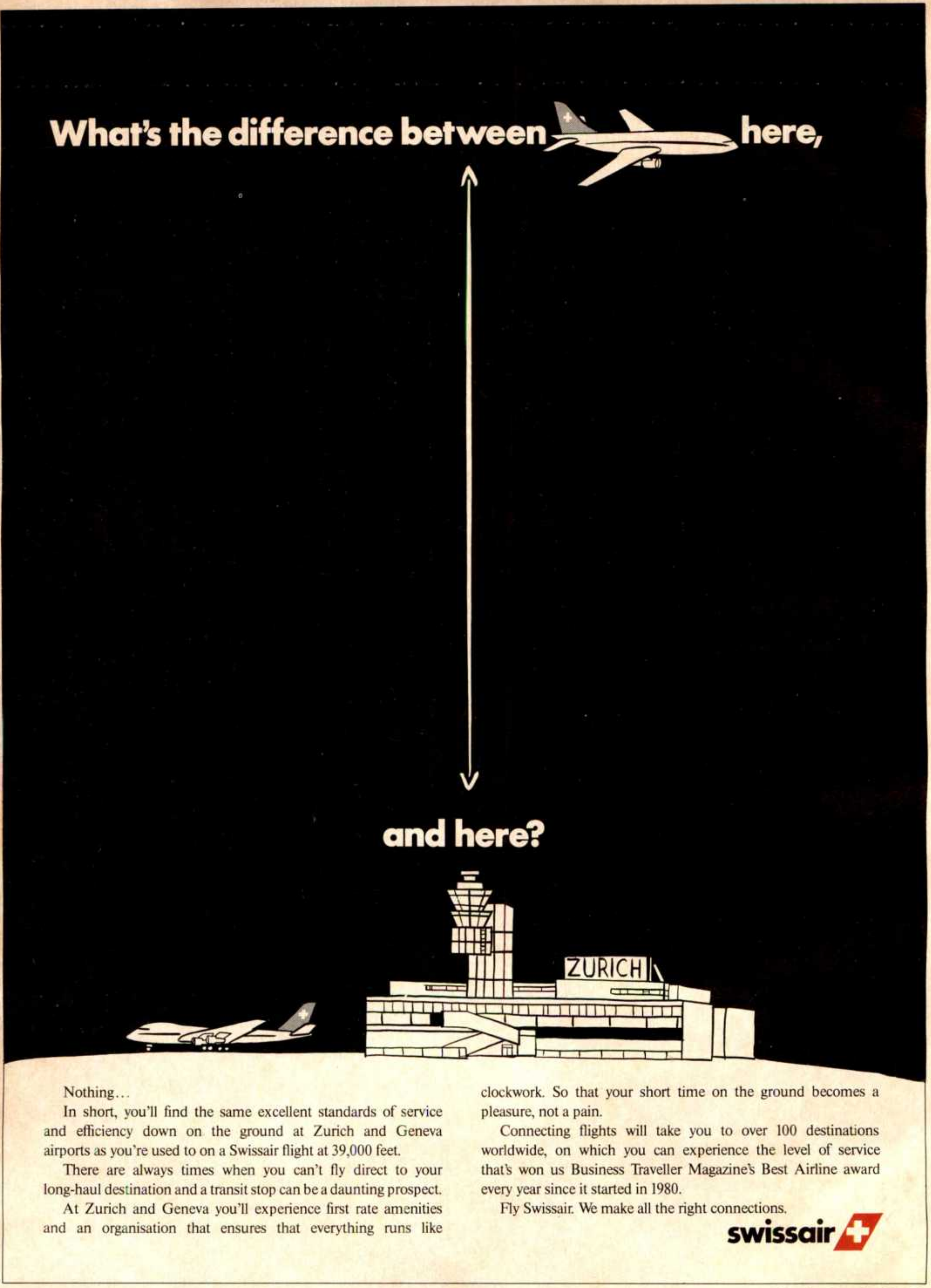
Matters arising are brief comments (up to 500 words) on articles and letters recently published in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted.

Manuscripts may be submitted either to London or Washington. Manuscripts should be typed (double spacing) on one side of the paper only. Four copies are required, each accompanied by copies of lettered artwork. No title should exceed 80 characters in length. Reference lists, figure legends, etc, should be on separate sheets, all of which should be numbered. Abbreviations, symbols, units, etc, should be identified on one copy of the manuscript at their first appearance.

References should appear sequentially indicated by superscripts, in the text and should be abbreviated according to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963-65). The first and last page numbers of each reference should be cited. Reference to books should clearly indicate the publisher and the date and place of publication. Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text.

Each piece of artwork should be clearly marked with the author's name and the figure number. Original artwork should be unlettered. Suggestions for cover illustrations are welcome. Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published.

Important: Manuscripts of proofs sent by air courier should be declared as "manuscripts" and "value $5" to prevent the imposition of Customs duty and Value Added Tax in the United Kingdom.


What's the difference between here,
and here?
ZURICH

Nothing…

In short, you'll find the same excellent standards of service and efficiency down on the ground at Zurich and Geneva airports as you're used to on a Swissair flight at 39,000 feet.

There are always times when you can't fly direct to your long-haul destination and a transit stop can be a daunting prospect.

At Zurich and Geneva you'll experience first rate amenities and an organisation that ensures that everything runs like clockwork. So that your short time on the ground becomes a pleasure, not a pain.

Connecting flights will take you to over 100 destinations worldwide, on which you can experience the level of service that's won us Business Traveller Magazine's Best Airline award every year since it started in 1980.

Fly Swissair. We make all the right connections.

swissair

Reader Service No.58

nature

From a Catalan atlas of 1375 by Abraham Crosques. This week we publish a review of science in Spain and Portugal. Photo: J Freeman/Fotomas.

OPINION

Arms, hostages and petroleum 289
Insiders outlawed 289
Teachers in travail 290

NEWS

Europe's framework research 291
US universities appear in good shape
UK research appointments
Magnetic levitation 292
Soviet reservations on IAEA
SERC scoops pool
US polar research 293
First leptonic collisions at Japan's KEK
Optical interferometry
Environmental databases 294
AIDS
Japanese volcano 295
Biotechnology insured
CNRS 296
Soviet academy
Networking the begging-bowl
Technology database 297
Information technology

CORRESPONDENCE

Singapore appeal/Fluoridation/UK performance/SDI boycott 298

NEWS AND VIEWS

How to make antimatter last 299
Contemplating life without sex
John Maynard Smith 300
Materials science: Nuts to tough ceramics
Paul Calvert 302
Semiconductors: an optical Stark effect observed in gallium arsenide
John Ryan 303
Plant pathology: host antifungal agents
Richard A Dixon 303
Problems with AIDS vaccines
Peter Newmark 304
A simple solution to the stability of the double helix?
Maxim Frank-Kamenetskii 305
Orbital evolution of comets
Julia Heisler 306

SCIENTIFIC CORRESPONDENCE

Role of HTLV III/LAV envelope protein
U Bicker 307
What is an extended probability?
W Muckenheim 307
Migration and hominid bipedalism
A R E Sinclair, M D Leakey & M Norton-Griffiths 307
^{131}I in milk and rain after Chernobyl
K Nishizawa, K Takata, N Hamada, Y Ogata, S Kojima, O Yamashita, M Ohshima & Y Kayama 308

BOOK REVIEWS

Unnatural Selection? Coming to Terms with the New Genetics *by E Yoxen*
D J Weatherall 309
The Rhinoceros from Dürer to Stubbs, 1515–1799 *by T H Clarke*
Martin Kemp 310
Evoked Potentials *R Q Cracco & I Bodis-Wollner, eds*
D Regan 310
The Biotechnological Challenge *S Jacobsson, A Jamison & H Rothman, eds*
Ronald Coleman 311
Fundamental Virology *B N Fields & D M Knipe, eds*
D H Watson 312
Trillion Year Spree: The History of Science Fiction *by B W Aldiss with D Wingrove*
John Treherne 312

SCIENCE IN IBERIA

A renaissance in the making 313
Spain: political change 314
Research structures under reform 315
World-beating molecular biology 319
The universities: problems overcome 320
Fighting swine fever virus 323
Boosting research in industry 324
Rebuilding the Basque Country 326
Portugal: politics and research 329
Remembering tropical colonies 331
Universities and industry 332

REVIEW ARTICLE

Superfluid ^{3}He in rotation
V P Mineev, M M Salomaa & O V Lounasmaa 333

ARTICLES

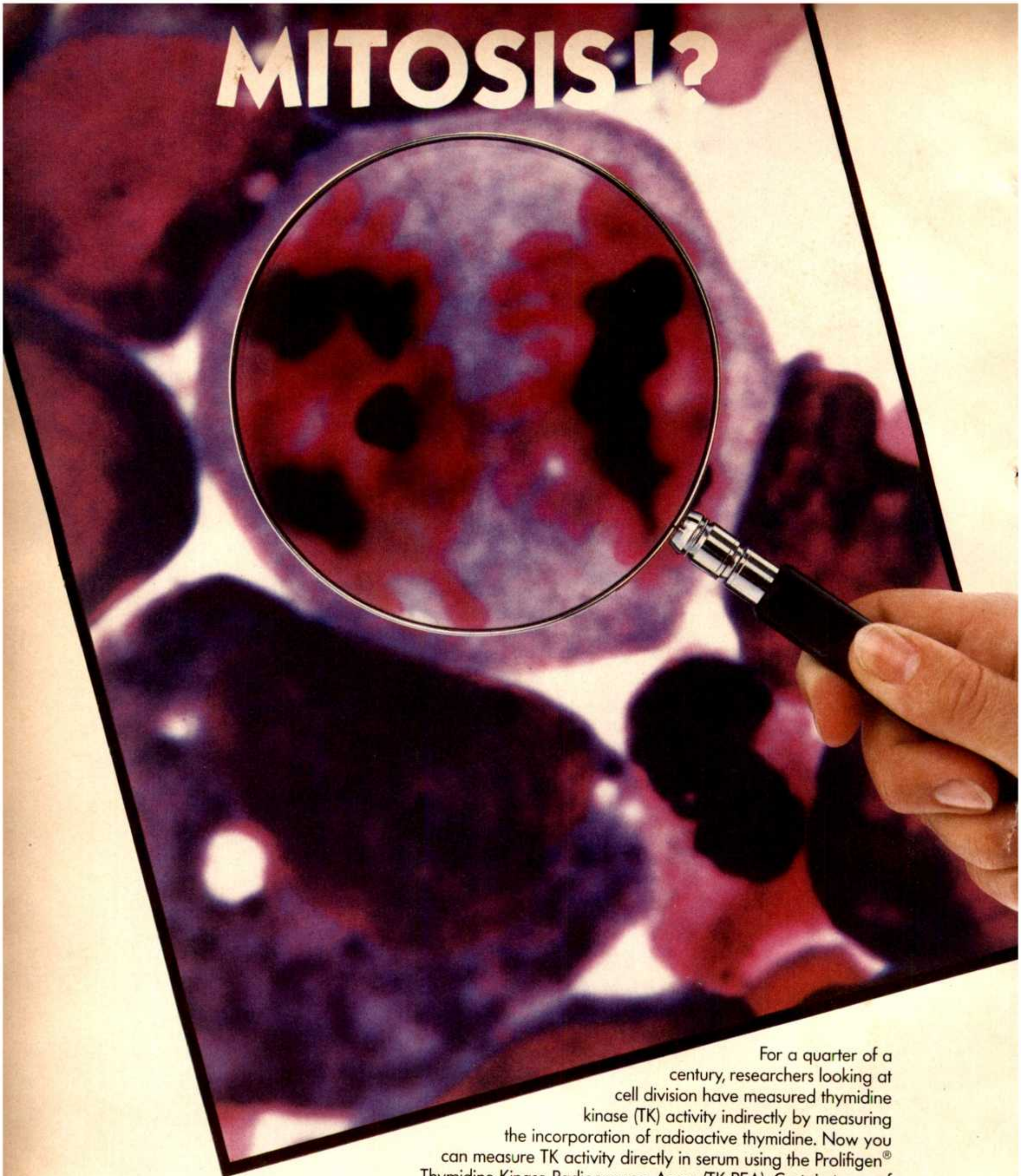
Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes
C W Pikielny, B C Rymond & M Rosbash 341

LETTERS TO NATURE

An accretion event in the Seyfert galaxy NGC 5548
B M Peterson & G J Ferland 345
Large-amplitude semidiurnal temperature variations in the polar mesopause: evidence of a pseudotide
R L Walterscheid, G G Sivjee, G Schubert & R M Hamwey 347
Perturbation of the background radiation by a moving gravitational lens
L I Gurvits & I G Mitrofanov 349
The near-parabolic flux and the origin of short-period comets
M E Bailey 350
Inferences on the evidence for radioactive ^{53}Mn in the early Solar System
T Lee 352
EISCAT observations of the E-region semidiurnal tide
T S Virdi, G O L Jones & P J S Williams 354
Long-range transport and deposition of acidic nitrogen in north-west Europe
R G Derwent & K Nodop 356
***Baryonyx*, a remarkable new theropod dinosaur**
A J Charig & A C Milner 359
Functional architecture of cortex revealed by optical imaging of intrinsic signals
A Grinvald, E Lieke, R D Frostig, C D Gilbert & T N Wiesel 361

Contents continued overleaf

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Journals Ltd and includes the Annual Index (mailed in February). Annual subscription for USA and Canada US $250. USA and Canadian orders to: *Nature*, Subscription Dept, PO Box 1501, Neptune, NJ 07753, USA. Other orders to *Nature*, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Second class postage paid at New York, NY 10012 and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $1.00 a copy plus $0.10 a page is paid directly to CCC, 21 Congress Street, Salem, Massachusetts 01970, USA. Identification code for *Nature*: 0028-0836/86 $1.00 + $0.10. US Postmaster send address changes to: *Nature*, 65 Bleecker Street, New York, NY 10012. Published in Japan by Macmillan Shuppan K.K., Eikow Building, 1-10-9 Hongo, Bunkyo-ku, Tokyo, Japan. © 1986 Macmillan Journals Ltd.



For a quarter of a century, researchers looking at cell division have measured thymidine kinase (TK) activity indirectly by measuring the incorporation of radioactive thymidine. Now you can measure TK activity directly in serum using the Prolifigen® Thymidine Kinase Radioenzyme Assay (TK-REA). Certain types of malignant disease are characterised by a particularly rapid cell proliferation; increased proliferation means an increased level of TK. Serum TK measurements have a proven clinical use for prognosis and follow-up in such diseases, particularly lymphomas, leukaemias and small-cell lung cancer. Contact Sangtec today and measure TK the easy way.

TK-REA is the way to measure Thymidine Kinase.

Sangtec MEDICAL

In U.S.A. for research purposes only

Box 20045, S-161 20 Bromma, Sweden Tel. Int + 46 8 98 83 30

Reader Service No.47

Sparre Medical Information, Sweden

CONTENTS

SCIENCE IN IBERIA

Spain and Portugal have recently emerged from social and political upheavals: how has the practice of science been affected?

In this special issue (27.11.86) Nature publishes a survey by Alun Anderson of new developments in both countries, in which he comments on key issues affecting Iberian scientists:

✻ What will Spain's new "law of science" achieve?

✻ Can Iberia reach the technological level needed to compete in the European Community?

✻ What have Iberia's researchers gained from their first decade of democracy?

Ordering information and prices as for back issues below.

Annual Subscription Prices

UK		£104
Rest of Europe	Airspeed	£125
USA & Canada		US$250
Australia, NZ & S. Africa	Airspeed	£150
	Airmail	£190
India	Airspeed	£120
	Airmail	£190
Japan	Airspeed	Y90000
Rest of World	Surface	£120
	Airmail	£180

Orders (with remittance) to:

USA & Canada	UK & Rest of World
Nature	Nature
Subscription Dept	Circulation Dept
PO Box 1501	Brunel Road
Neptune	Basingstoke
NJ 07753	Hants RG21 2XS, UK
USA	Tel: 0256 29242

(The addresses of Nature's editorial offices are shown facing the first editorial page)

Japanese subscription enquiries to:
Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan
Tel: (03) 292 3755

Personal subscription rates: These are available in some countries to subscribers paying by personal cheque or credit card. Details from:

USA & Canada	UK & Europe
Nature	Felicity Parker
65 Bleecker Street	Nature
New York	4 Little Essex Street
NY 10012, USA	London WC2R 3LF, UK
Tel: (212) 477-9600	Tel: 01-836 6633

Back issues: UK, £2.50; USA & Canada, US$6.00 (surface), US$9.00 (air); Rest of World, £3.00 (surface), £4.00 (air)

Binders
Single binders: UK, £5.25; USA & Canada $11.50; Rest of World, £7.75. Set of four: UK, £15.00; USA & Canada $32.00; Rest of World, £22.00

Annual indexes (1971-1985)
UK, £5.00 each; Rest of World, $10.00

Nature First Issue Facsimile
UK, £2.00; Rest of World (surface), $3.00; (air) $4.00

Nature in microform
Write to University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA

Control of malarial invasion by phosphorylation of the host cell membrane cytoskeleton
K Rangachari, A Dluzewski, R J M Wilson & W B Gratzer 364

Plant chitinases are potent inhibitors of fungal growth
A Schlumbaum, F Mauch, U Vögeli & T Boller 365

Sulphonolipids are molecular determinants of gliding motility
D R Abbanat, E R Leadbetter, W Godchaux III & E Escher 367

Ion channels in human neutrophils activated by a rise in free cytosolic calcium concentration
V von Tscharner, B Prod'hom, M Baggiolini & H Reuter 369

A complementary DNA clone for a macrophage-lymphocyte Fc receptor
V A Lewis, T Koch, H Plutner & I Mellman 372

Loss of mouse fibroblast cell response to phorbol esters restored by microinjected protein kinase C
G Pasti, J-C Lacal, B S Warren, S A Aaronson & P M Blumberg 375

The v-*fms* oncogene induces factor independence and tumorigenicity in CSF-1 dependent macrophage cell line
E F Wheeler, C W Rettenmier, A T Look & C J Sherr 377

Recurrent mutations in haemophilia A give evidence for CpG mutation hotspots
H Youssoufian, H H Kazazian Jr, D G Phillips, S Aronis, G Tsiftis, V A Brown & S E Antonarakis 380

The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein
M Z Papiz, L Sawyer, E E Eliopoulos, A C T North, J B C Findlay, R Sivaprasadarao, T A Jones, M E Newcomer & P J Kraulis 383

DNA conformation is determined by economics in the hydration of phosphate groups
W Saenger, W N Hunter & O Kennard 385

A protein on *Plasmodium falciparum*-infected erythrocytes functions as a transferring receptor
M H Rodriguez & M Jungery 388

MATTERS ARISING

Arc minute gravitational lenses and cosmic strings
D P Bennett & A Stebbins
Reply: P Paczyński 392

'Capture' of stereopsis by illusory contours
K Prazdny
Reply: V S Ramachandran 393

Is the Gulf of Tadjoura a platelet boundary or complex rift zone?
M Boucarut & M Clin 394

NATURE CLASSIFIED

Professional appointments — Research posts — Studentships — Fellowships — Conferences — Courses — Seminars — Symposia: Back Pages

Next week in *Nature*:

- Creationism versus evolution
- More on Halley
- New computing algorithm
- Natural/unnatural lake acidification
- Pollution and seals
- Man in New Guinea 40,000 years ago
- Insulin secretion inhibited
- New way to atherosclerosis drugs?
- Resistance to chemotherapy
- Immortal B lymphocytes

GUIDE TO AUTHORS

Authors should be aware of the diversity of *Nature*'s readership and should strive to be as widely understood as possible.

Review articles should be accessible to the whole readership. Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable).

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding. The text should not exceed 3,000 words and six displayed items (figures plus tables). The article should include an italic heading of about 50 words.

Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items. The first paragraph (not exceeding 150 words) should say what the letter is about, why the study it reports was undertaken and what the conclusions are.

Matters arising are brief comments (up to 500 words) on articles and letters recently published in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted.

Manuscripts may be submitted either to London or Washington. Manuscripts should be typed (double spacing) on one side of the paper only. Four copies are required, each accompanied by copies of lettered artwork. No title should exceed 80 characters in length. Reference lists, figure legends, etc, should be on separate sheets, all of which should be numbered. Abbreviations, symbols, units, etc, should be identified on one copy of the manuscript at their first appearance.

References should appear sequentially indicated by superscripts, in the text and should be abbreviated according to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963-65). The first and last page numbers of each reference should be cited. Reference to books should clearly indicate the publisher and the date and place of publication. Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text.

Each piece of artwork should be clearly marked with the author's name and the figure number. Original artwork should be unlettered. Suggestions for cover illustrations are welcome. Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published.

Important: Manuscripts of proofs sent by air courier should be declared as "manuscripts" and "value $5" to prevent the imposition of Customs duty and Value Added Tax in the United Kingdom.

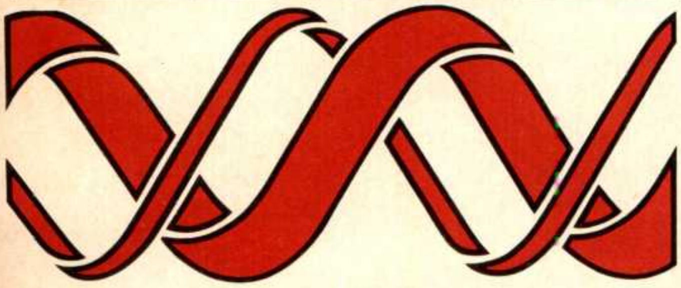
Requests for permission to reproduce material from Nature *should be accompanied by a self-addressed (and, in the case of the United Kingdom and United States, stamped) envelope.*

ANNOUNCIN

BIO/TECHNOLOGY NEW

JANUARY 2

• GENE REGULATION • RECEPTO



CONFERENCE PROGRAM

Plenary Address
George Poste
Smith Kline & French Research Laboratories

DAY 1 • Regulation of Gene Expression

Satellite Phage P4
Richard Calendar, University of California, Berkeley

Soybean Seed Protein in Transformed Plants
Robert Goldberg, University of California, Los Angeles

The Immunoglobulin Gene Superfamily
Leroy Hood, California Institute of Technology

Transcriptional Activation
Martin Rosenberg, Smith Kline & French Research Laboratories

Organ-specific Gene Expression in Chimeric Plants
Marc Van Montagu, Rijsuniversiteit Gent

Alpha Interferon
Charles Weissmann, University of Zurich

DAY 2 • Receptor Biology: Towards New Pharmaceuticals

Receptor Regulation and Hypercholesterolemia
Michael Brown, University of Texas Southwest Medical School

Antireceptor Antibodies
Mark Greene, University of Pennsylvania

Epidermal Growth Factor Receptor Gene
Ira Pastan, National Cancer Institute

The T-Cell Receptor
Susumu Tonegawa, Massachusetts Institute of Technology

Growth Factor Receptor Tyrosine Kinases
Axel Ullrich, Genentech, Inc.

The Interleukin-2 Receptor
Thomas Waldmann, National Cancer Institute

DAY 3 • Redesigning Proteins

After the Cloning is Over
Charles Craik, University of California, San Francisco

A Biophysical Approach to Protein Secretion
Lila Gierasch, University of Delaware

Recovering and Refolding Active Oligomeric Proteins
Rainer Jaenicke, Universitat Regensburg

Genetic Indentification of Protein Folding Pathways
Jonathan King, Massachusetts Institute of Technology

Modifying Enzymatic Activity and Stability
Gregory Petsko, Massachusetts Institute of Technology

Disulfide Bonds and Protein Stability
Ronald Wetzel, Genentech, Inc.

The second Bio/Technology Conference will tackle the many scientific issues in commercial protein production. Understanding receptor biology is essential to developing everything from vaccines to cancer therapeutics to heart drugs. Increasing dependence on commercial production in mammalian cells makes it critically important to understand the development of eukaryotic gene regulation. And recovering and refolding complex proteins into bioactive forms is the key step in making usable products.

Next January in New Orleans, leading scientists from around the world will come together to discuss their most recent research, and to explore three key issues essential to the rapid development of biotechnology.

- **Gene regulation and expression**
- **Receptor genetics and function**
- **Structure/function relationships in proteins.**

You will be updated on work that is leading to new pharmaceuticals. We'll discuss the protein folding problem. Rearrangement of the DNA template. Details of the molecular structure and function of receptors such as ligand processing and drug effects. Transactivation of viral and mammalian genes.

This conference is an exceptional opportunity to hear from an outstanding group of speakers, and to exchange ideas and information with them — as well as with other conference delegates.

New Products on View

In conjunction with the conference, there will be an exhibition of the latest scientific products and services offered by leading research labs, pharmaceutical firms, publishers and many others. Representatives will be on hand to discuss their products and to answer your questions.

Submit Your Work for Poster Sessions

There will be a limited number of opportunities for poster presentations, and we invite scientists who are interested to submit abstracts of relevant work for consideration. Mail abstracts to:

Research Editor
Bio/Technology
65 Bleecker Street
New York, NY 10012

Registration will be limited, so be sure to register early.
BIO/TECHNOLOGY subscribers save an additional $50.

nature

The huge hydrogen cloud surrounding the nucleus of comet Halley, as seen in the ultraviolet (Lyman-α). The direction of the Sun is to the bottom of the image. The outermost yellow ring is about 10 million kilometres across its long axis. The false colour coding distinguishes regions of different brightness and is multivalued to enhance contrast. The image was taken from a sounding rocket on 13 March. Studies of Halley from rocket, aircraft and satellite are reported on pages 433–444.

OPINION

Europe runs before it walks 395

Licensing the spooks 396

Action against AIDS

NEWS

French universities 397

Hermes to win approval soon?

Chemical weapons 398

National Science Foundation

Telecommunications 399

Hepatitis B

India creates biosphere reserve

US biotechnology 400

Information technology

Alaskan environment

Laboratory fraud 401

Science teachers 402

Biotechnology companies

Eastern Europe 403

US defence

CORRESPONDENCE

Nuclear risk assessment/Avalanche dowsing/Wrong dose 404

NEWS AND VIEWS

Images in and of the brain
Jennifer Altman 405

Mathematics: One hundred per cent proof
Ian Stewart 406

Progress in understanding multidrug resistance
George R Stark 407

Hard-wired local triggering of intestinal enzyme expression
Jared M Diamond 408

New discovery of waisted axes 409

Obituary: Albert Szent-Györgyi (1893–1986)
John T Edsall 409

Toxins and grazing resistance
Peter D Moore 410

Galactic evolution: Do globular clusters belong?
Virginia Trimble 411

New hopes for the treatment of coronary heart disease
Gilbert Thompson 412

Big, repulsive boson stars
Robert Walgate 413

Plant molecular biology: tailoring crop improvement
Chris Lamb & Leona Fitzmaurice 414

SCIENTIFIC CORRESPONDENCE

BSF-2 is not just a differentiation factor
A Billiau 415

Plant tumour induction
C H Shaw, A M Ashby & M D Watson 415

An alternative view of the origin of life
N G Mehta 415

The kingdoms of organisms
T Cavalier-Smith 416

How many reactor accidents?
A W F Edwards 417

Stunning whales
D C Jolly 418

Body temperature and the specific heat of water
W A Calder III, J D Dunitz & S A Benner 418

BOOK REVIEWS

Supersymmetry and its Applications: Superstrings, Anomalies and Supergravity *G W Gibbons, S W Hawking and P K Townsend, eds*
Michael Atiyah 419

Spiders: Webs, Behavior, and Evolution *W A Shear, ed.*
J L Cloudsley-Thompson 420

The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics *P C W Davies and J R Brown, eds*
Michael Redhead 420

Environmental Magnetism *by R Thompson and F Oldfield*
Subir K Banerjee 421

Nicholas Copernicus: Complete Works. Vol. III *P Czartoryski, ed., translation and commentary by E Rosen with E Hilfstein*
Robert S Westman 422

COMMENTARY

US Supreme Court to review Louisiana appeal
T H Jukes 423

ARTICLES

A thyromimetic that decreases plasma cholesterol levels without increasing cardiac activity
A H Underwood, J C Emmett, D Ellis, S B Flynn, P D Leeson, G M Benson, R Novelli, N J Pearce & V P Shah 425

The *Tetrahymena* ribozyme acts like an RNA restriction endonuclease
A J Zaug, M D Been & T R Cech 429

LETTERS TO NATURE

Is CO_2 responsible for the outbursts of comet Halley?
P D Feldman, M F A'Hearn, M C Festou, L A McFadden, H A Weaver & T N Woods 433

Rocket ultraviolet spectroscopy of comet Halley and abundance of carbon monoxide and carbon
T N Woods, P D Feldman, K F Dymond & D J Sahnow 436

Far-ultraviolet spectral images of comet Halley from sounding rockets
R P McCoy, C B Opal & G R Carruthers 439

Post-perihelion observations of water in comet Halley
H A Weaver, M J Mumma, H P Larson & D S Davis 441

Identification of [Fe III] in the solar ultraviolet spectrum
C Jordan, J-D F Bartoe, G E Brueckner & K P Dere 444

A hierarchical O ($N \log N$) force-calculation algorithm
J Barnes & P Hut 446

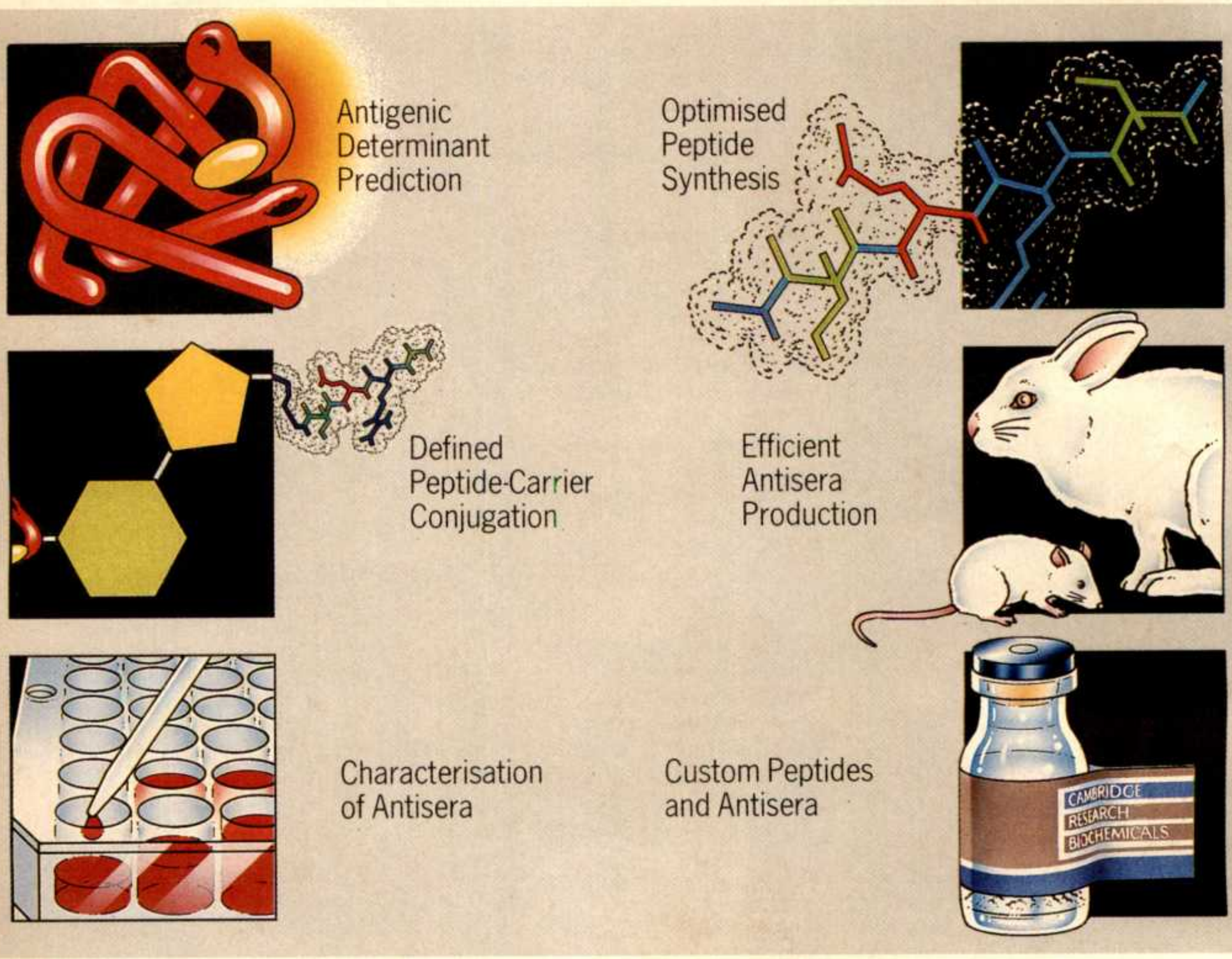
The 400-km seismic discontinuity and the proportion of olivine in the Earth's upper mantle
C R Bina & B J Wood 449

Natural and anthropogenic causes of lake acidification in Nova Scotia
E Gorham, J K Underwood, F B Martin & J G Ogden III 451

Contents continued overleaf

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Journals Ltd and includes the Annual Index (mailed in February). Annual subscription for USA and Canada US $250. USA and Canadian orders to: *Nature*, Subscription Dept, PO Box 1501, Neptune, NJ 07753, USA. Other orders to *Nature*, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Second class postage paid at New York, NY 10012 and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $1.00 a copy plus $0.10 a page is paid directly to CCC, 21 Congress Street, Salem, Massachusetts 01970, USA. Identification code for *Nature*: 0028-0836/86 $1.00 + $0.10. US Postmaster send address changes to: *Nature*, 65 Bleecker Street, New York, NY 10012. Published in Japan by Macmillan Shuppan K.K., Eikow Building, 1-10-9 Hongo, Bunkyo-ku, Tokyo, Japan. © 1986 Macmillan Journals Ltd.

Custom Peptides and Antisera...



...the Complete Service.

An expanded custom service for the rapid production of synthetic peptides and antipeptide antibodies is now available from the specialists in peptide research, Cambridge Research Biochemicals.

Developed to meet all your project needs, this confidential service includes hydropathic analysis to identify potential antigenic determinants and the efficient Fmoc-polyamide synthesis of selected peptides which are subsequently fully characterised.

Employing defined peptide-carrier conjugation and optimised immunisation protocols together with the assay of sera for anti*peptide* activity, the Cambridge Research Biochemicals custom service provides you with a unique product of the required pre-determined specificity.

Contact Dr Paul Sheppard, Contract Research Director, now for further details of this exciting new service or simply send for a free copy of the new 'Synthetic Peptides For Molecular Biology' brochure and discover how we produce the best at Cambridge.

CAMBRIDGE
RESEARCH
BIOCHEMICALS

UK. Cambridge Research Biochemicals Ltd., Button End, Harston, Cambridge, CB2 5NX, England. Tel: Cambridge (0223) 871674. Telex: 817694 CRB LTD G. Fax: 0223 872381
USA. Cambridge Research Biochemicals Inc., PO Box 58, 2005 Park Street, Atlantic Beach, New York 11509, USA. Toll free Tel: (800) 327 0125. Telex: 971398. Fax: 516 239 2782

Reader Service No.166

SCIENCE IN IBERIA

Spain and Portugal have recently emerged from social and political upheavals: how has the practice of science been affected?

In this special issue (27.11.86) Nature publishes a survey by Alun Anderson of new developments in both countries, in which he comments on key issues affecting Iberian scientists:

- What will Spain's new "law of science" achieve?
- Can Iberia reach the technological level needed to compete in the European Community?
- What have Iberia's researchers gained from their first decade of democracy?

Ordering information and prices as for back issues below.

Annual Subscription Prices

UK		£104
Rest of Europe	Airspeed	£125
USA & Canada		US$250
Australia, NZ & S. Africa	Airspeed	£150
	Airmail	£190
India	Airspeed	£120
	Airmail	£190
Japan	Airspeed	Y90000
Rest of World	Surface	£120
	Airmail	£180

Orders (with remittance) to:

USA & Canada	**UK & Rest of World**
Nature	Nature
Subscription Dept	Circulation Dept
PO Box 1501	Brunel Road
Neptune	Basingstoke
NJ 07753	Hants RG21 2XS, UK
USA	Tel: 0256 29242

(The addresses of Nature's editorial offices are shown facing the first editorial page)

Japanese subscription enquiries to:
Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan
Tel: (03) 292 3755

Personal subscription rates: These are available in some countries to subscribers paying by personal cheque or credit card. Details from:

USA & Canada	**UK & Europe**
Nature	Felicity Parker
65 Bleecker Street	Nature
New York	4 Little Essex Street
NY 10012, USA	London WC2R 3LF, UK
Tel: (212) 477-9600	Tel: 01-836 6633

Back issues: UK, £2.50; USA & Canada, US$6.00 (surface), US$9.00 (air); Rest of World, £3.00 (surface), £4.00 (air)

Binders
Single binders: UK, £5.25; USA & Canada $11.50; Rest of World, £7.75. Set of four: UK, £15.00; USA & Canada $32.00; Rest of World, £22.00

Annual indexes (1971-1985)
UK, £5.00 each; Rest of World, $10.00

Nature First Issue Facsimile
UK, £2.00; Rest of World (surface), $3.00; (air) $4.00

Nature in microform
Write to University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA

CONTENTS

A 40,000 year-old human occupation site at Huon Peninsular, Papua New Guinea
L Groube, J Chappell, J Muke & D Price 453

Reproductive failure in common seals feeding on fish from polluted coastal waters
P J H Reijnders 456

A common mammalian plan of accessory optic system organization revealed in all primates
H M Cooper & M Magnin 457

A cell surface molecule distributed in a dorsoventral gradient in the perinatal rat retina
M Constantine-Paton, A S Blum, R Mendez-Otero & C J Barnstable 459

Differentiation state-dependent surface mobilities of two forms of the neural cell adhesion molecule
G E Pollerberg, M Schachner & J Davoust 462

Endfeet of retinal glial cells have higher densities of ion channels that mediate K^+ buffering
H Brew, P T A Gray, P Mobbs & D Attwell 466

Cyclic GMP is involved in the excitation of invertebrate photoreceptors
E C Johnson, P R Robinson & J E Lisman 468

Characterization of single pacemaker channels in cardiac sino-atrial node cells
D DiFrancesco 470

Atrial natriuretic peptide causes pre-glomerular vasodilatation and post-glomerular vasoconstriction in rat kidney
M Marin-Grez, J T Fleming M Steinhausen 473

Pancreastatin, a novel pancreatic peptide that inhibits insulin secretion
K Tatemoto, S Efendić, V Mutt, G Makk, G J Feistner & J D Barchas 476

Glucocorticoid receptors lacking hormone-binding activity are bound in nuclei of ATP-depleted cells
D B Mendel, J E Bodwell F A Munck 478

A new subunit of the human T-cell antigen receptor complex
A M Weissman, L E Samelson R D Klausner 480

Athymic mice express a high level of functional γ-chain but greatly reduced levels of α- and β-chain T-cell receptor messages
Y Yoshikai, M D Reis & T W Mak 482

Homology between P-glycoprotein and a bacterial haemolysin transport protein suggests a model for multidrug resistance
J H Gerlach, J A Endicott, P F Juranka, G Henderson, F Sarangi, K L Deuchars & V Ling 485

Mature murine B lymphocytes immortalized by Kirsten sarcoma virus
A H Lichtman, D S Reynolds, D V Faller & A K Abbas 489

Energy conversion catalysis using semiconducting transition metal cluster compounds: Erratum
N A Vante & H Tributsch 491

MATTERS ARISING

Stratigraphic record of earthquake incidence
A G Fischer 492

Primary structure of the histidine-rich protein of *Plasmodium lophurae*
D O Irving & G A M Cross 492

PRODUCT REVIEW

Quantitative receptor autoradiography
M D Hall, A P Davenport & C R Clark 493

International exhibits 494

EMPLOYMENT

A strategy for higher education
R Pearson 498

NATURE CLASSIFIED

Professional appointments — Research posts — Studentships — Fellowships — Conferences — Courses — Seminars — Symposia: Back Pages

MISCELLANY

100 years ago 410

Next week in *Nature*:

- Music of digital computers
- Cosmic microwave background anisotropy.
- Gradients and pattern formation
- New low-dimensional metal
- Salt-based radioactive waste storage
- Origin of ratites
- Antigens of the AIDS virus
- Antigen recognition by T-lymphocytes

BIOSYSTEMS UPDATE

For People who make Peptides

Here are five reasons why.

1 It's the Truly Efficient Way to Make Peptides

The Model 430A Peptide Synthesizer doesn't waste time, amino acids, or reagents. It consistently produces high yields of quality peptide which reduces post-synthesis purification and analysis significantly. You get more peptide for your effort *and* for your dollar.

2 It's More Flexible, Easier to Operate

The 430A meets the needs of every user whether new to peptide synthesis or very experienced. Software is written in plain English using familiar terminology. All control is implemented by simply touching the screen. Standard synthesis cycles and functions can be selected or new ones designed quickly and easily.

3 Its Chemistry is Proven and Unique

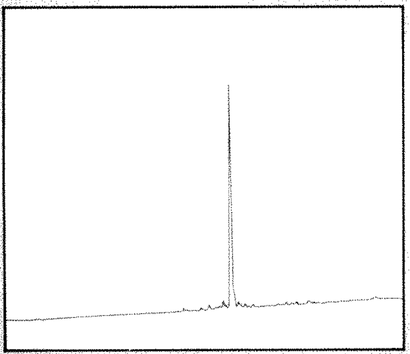
High stepwise yields, high purity and reproducibility are routine. Each step of the chemistry is specific to the requirements of the various amino acids. Coupling failures are eliminated or greatly reduced resulting in fewer double couplings. Both *t*-Boc and FMOC chemistries can be accommodated.

4 It Produces Superb Short Peptides

Step Yield (% Coupled) Standard ABI *t*-Boc cycles, version 1.0

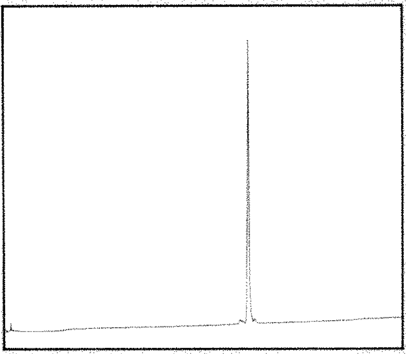
PEPTIDE SEQUENCE	LEU	HIS	PHE	PRO	HIS	ILE	TYR	VAL	ARG	ASP	AVG. STP. YIELD %
Ninhydrin Analysis	•	96.7	99.7	99.7	†99.6	98.9	99.5	98.3	99.5	99.4	**99.0**
Preview Sequence Analysis	•	•	•	99.5	99.2	94.9	96.6	96.5	97.1	97.9	**97.4**
RESIDUE NUMBER	10	9	8	7	6	5	4	3	2	1	

†Ninhydrin values for amino acids coupled to proline are not quantitative.



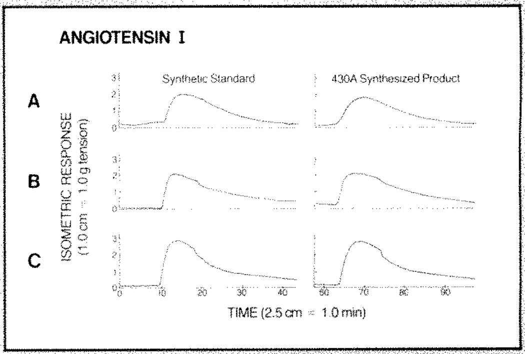
Angiotensin I, Crude

mobile phase	0.1% TFA/CH_3CN
gradient	0 to 60% in 60 minutes
flow rate	1.0 mL/min
detector	216 nm
column	Brownlee, C-8, 300Å, 10μm 4.6 mm x 250 mm



Angiotensin I, HPLC Purified

mobile phase	0.1% TFA/CH_3CN
gradient	0 to ≥ 60% in 60 minutes
flow rate	1.0 mL/min
detector	216 nm
column	Brownlee, C-8, 300Å, 10μm 4.6 mm x 250 mm



Panels A, B and C are isometric responses of two different Ile^5-Angiotensin I—a synthetic standard and the 430A-produced, purified peptide—on three separate rabbit aortic strips. The responses produced by the Applied Biosystems Ile^5-Angiotensin I were not significantly different than those of the standard. Sample concentrations were: in panel A, 564 ng/mL; in panels B and C, 500 ng/mL.

These response data were furnished by Drs. M. C. Khosla and A. Husein of the Cleveland Clinic Foundation.

Copyright 1986, Applied Biosystems, Inc.

nature

Reflected light micrograph of a section of rock salt from the Asse Mine (Germany). Rock salt is important as a material for use in radioactive waste depositories and for its association with oil traps. The effect of water on the low-stress rheology of rock salt is discussed on pages 554–557 of this issue and also in a News and Views article on page 515.

OPINION

Another nuclear power station? 499
Shades of 1968? 500

NEWS

After the riots the politics 501
European summit 502
Independence for UK space centre?
Super-collider still fighting for funds
Amazon rainforests 503
Space shuttle
Climatic prediction 504
Biotechnology patents
Bhopal 505
Space Station cooperation
Monoclonal patent 506
AIDS research queried
New hepatitis-B vaccine launched
Japanese technology 507
Indian resources
UK – Dutch telescope

CORRESPONDENCE

Italian research/Maoris/ Panselectionism 508

NEWS AND VIEWS

When firsts come only second 509
Solid-state chemistry: Novel molecular metals
Martin R Bryce 510
Pattern formation: Form and diffusion
Alfonso Martinez-Arias & Philip Ingham 510
Nuclear fusion comes closer
R S Pease 512
How many species are there?
Robert M May 514
Halite rheology: Salt with a pinch of water
Malcolm K Jenyon 515
Phylogeny: The German 'ostrich' and the molecular clock
Michael E Howgate 516

SCIENTIFIC CORRESPONDENCE

The highest inhabitants in the world
J B West 517
Functions of leaf fall
R C Hardwick 517
Undirectional spin
H Bondi 517
Detecting Earth-like planets
B F Burke
Reply: R Angel, A Cheng & N Woolf 518

BOOK REVIEWS

The Rise of Statistical Thinking 1820–1900 *by T M Porter*
and
The History of Statistics: The Measurement of Uncertainty Before 1900 *by S M Stigler*
Ian Stewart 519
The Nordic Seas
B G Hurdle, ed.
W J Gould & D G Masson 520
Living Together: The Biology of Animal Parasitism
by W Trager
Sydney Cohen 521
The Riddle of the Dinosaur
by J N Wilford
Alan J Charig 521
God and the New Biology
by A Peacocke
William Paton 522

COMMENTARY

The music of digital computers
P Campbell 523

REVIEW ARTICLE

Cosmic microwave background anisotropy
N Kaiser & J Silk 529

ARTICLES

A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern
P M Macdonald & G Struhl 537

LETTERS TO NATURE

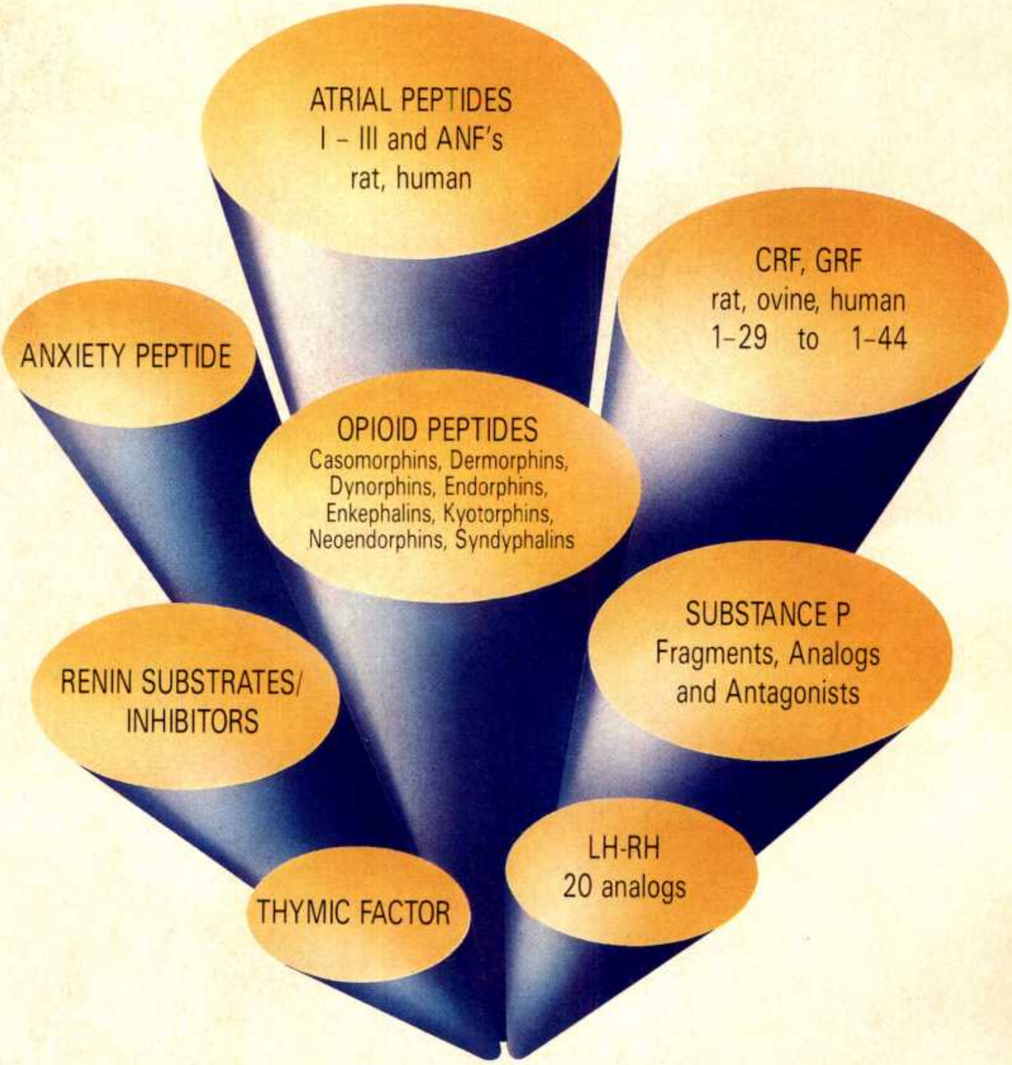
Infrared variability of the BL lacertae object OJ287 since its outburst in 1983
W K Gear, E I Robson & L M J Brown 546
A new low-dimensional metal, $Cs[Pd(S_2C_2(CN)_2)_2].0.5\ H_2O$
I D Parker, R H Friend, P I Clemenson & A E Underhill 547
Organ pipe radiant modes of periodic micromachined silicon surfaces
P J Hesketh, J N Zemel & B Gebhart 549
Evidence for a link between atmospheric thermonuclear detonations and nitric acid
G Holdsworth 551
Weakening of rock salt by water during long-term creep
J L Urai, C J Spiers, H J Zwart & G S Lister 554
Diagenesis of US Gulf Coast shales
K Pye, D H Krinsley & J H Burton 557
Do continental shelves export organic matter?
G T Rowe, S Smith, P Falkowski, T Whitledge, R Theroux, W Phoel & H Ducklow 559
Identification of novel widely distributed sedimentary acyclic sesterterpenoids
J N Robson & S J Rowland 561
Ostrich ancestors found in the Northern Hemisphere suggest new hypothesis of ratite origins
P Houde 563
Phospholipid vesicles as a model system for biomineralization
S Mann, J P Hannington & R J P Williams 565

Contents continued overleaf

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Journals Ltd and includes the Annual Index (mailed in February). Annual subscription for USA and Canada US $250. USA and Canadian orders to: *Nature*, Subscription Dept, PO Box 1501, Neptune, NJ 07753, USA. Other orders to *Nature*, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Second class postage paid at New York, NY 10012 and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $1.00 a copy plus $0.10 a page is paid directly to CCC, 21 Congress Street, Salem, Massachusetts 01970, USA. Identification code for *Nature*: 0028-0836/86 $1.00 + $0.10. US Postmaster send address changes to: *Nature*, 65 Bleecker Street, New York, NY 10012. Published in Japan by Macmillan Shuppan K.K., Eikow Building, 1-10-9 Hongo, Bunkyo-ku, Tokyo, Japan. © 1986 Macmillan Journals Ltd.

SWISS-BACHEM

INDUSTRIAL QUANTITIES OF
BIOLOGICALLY ACTIVE PEPTIDES
a selection of SWISS-BACHEM'S products



BACHEM AG HAUPTSTRASSE 144
CH – 4416 BUBENDORF, SWITZERLAND
TELEX 966 081 BACE CH
Reader Service No.88

CONTENTS

SCIENCE IN IBERIA

Spain and Portugal have recently emerged from social and political upheavals: how has the practice of science been affected?

In this special issue (27.11.86) Nature publishes a survey by Alun Anderson of new developments in both countries, in which he comments on key issues affecting Iberian scientists:

✻ What will Spain's new "law of science" achieve?

✻ Can Iberia reach the technological level needed to compete in the European Community?

✻ What have Iberia's researchers gained from their first decade of democracy?

Ordering information and prices as for back issues below.

Annual Subscription Prices

UK		£104
Rest of Europe	Airspeed	£125
USA & Canada		US$250
Australia, NZ & S. Africa	Airspeed	£150
	Airmail	£190
India	Airspeed	£120
	Airmail	£190
Japan	Airspeed	Y90000
Rest of World	Surface	£120
	Airmail	£180

Orders (with remittance) to:

USA & Canada	UK & Rest of World
Nature	Nature
Subscription Dept	Circulation Dept
PO Box 1501	Brunel Road
Neptune	Basingstoke
NJ 07753	Hants RG21 2XS, UK
USA	Tel: 0256 29242

(The addresses of Nature's editorial offices are shown facing the first editorial page)

Japanese subscription enquiries to:
Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan
Tel: (03) 292 3755

Personal subscription rates: These are available in some countries to subscribers paying by personal cheque or credit card. Details from:

USA & Canada	UK & Europe
Nature	Felicity Parker
65 Bleecker Street	Nature
New York	4 Little Essex Street
NY 10012, USA	London WC2R 3LF, UK
Tel: (212) 477-9600	Tel: 01-836 6633

Back issues: UK, £2.50; USA & Canada, US$6.00 (surface), US$9.00 (air); Rest of World, £3.00 (surface), £4.00 (air)

Binders
Single binders: UK, £5.25; USA & Canada $11.50; Rest of World, £7.75. Set of four: UK, £15.00; USA & Canada $32.00; Rest of World, £22.00

Annual indexes (1971-1985)
UK, £5.00 each; Rest of World, $10.00

Nature First Issue Facsimile
UK, £2.00; Rest of World (surface), $3.00; (air) $4.00

Nature in microform
Write to University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA

A thylakoid processing protease is required for complete maturation of the lumen protein plastocyanin
J Hageman, C Robinson, S Smeekens & P Weisbeek 567

Cholinergic neurones acquire adrenergic neurotransmitters when transplanted into an embryo
J N Coulombe & M Bronner-Fraser 569

Variable and conserved neutralization antigens of human immunodeficiency virus
R A Weiss, P R Clapham, J N Weber, A G Dalgleish, L A Lasky & P W Berman 572

Cytotoxic T lymphocytes recognize influenza haemagglutinin that lacks a signal sequence
A R M Townsend, J Bastin, K Gould & G G Brownlee 575

H-2-restricted cytolytic T cells specific for HLA can recognize a synthetic HLA peptide
J L Maryanski, P Pala, C Corradin, B R Jordan & J-C Cerottini 578

Lambda 5, a new light-chain-related locus selectively expressed in pre-B lymphocytes
N Sakaguchi & F Melchers 579

Long-range restriction map around the Duchenne muscular dystrophy gene
M Burmeister & H Lehrach 582

Recombination between immunoglobulin variable region segments is enhanced by transcription
T K Blackwell, M W Moore, G D Yancopoulos, H Suh, S Lutzker, E Selsing & F W Alt 585

Effect of mutations at the lariat branch acceptor site on β-globin pre- mRNA splicing *in vitro*
H Hornig, M Aebi & C Weissmann 589

The correct activation of *Antennapedia* and bithorax complex genes requires the *fushi tarazu* gene
P W Ingham & A Martinez-Arias 592

Does nuclear integration of mitochondrial sequences occur during senescence in *Podospora*?
F Koll 597

MATTERS ARISING

Thermal erosion by komatiites at Kambalda
N T Arndt 600

MISCELLANY

100 years ago 513

NATURE CLASSIFIED

Professional appointments — Research posts — Studentships — Fellowships — Conferences — Courses — Seminars — Symposia: Back Pages

Next week in *Nature*:

- Cyanogen jets in Halley
- Controlling exhaust gases
- Ozone-layer 'hole'
- Bacterial genetics 40 years on
- Function of *c-erb-A* oncogene
- Cloning AIDS HIV-2
- *Drosophila* gap genes
- Making proteases heat-stable

GUIDE TO AUTHORS

Authors should be aware of the diversity of ***Nature***'s readership and should strive to be as widely understood as possible.

Review articles should be accessible to the whole readership. Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable).

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding. The text should not exceed 3,000 words and six displayed items (figures plus tables). The article should include an italic heading of about 50 words.

Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items. The first paragraph (not exceeding 150 words) should say what the letter is about, why the study it reports was undertaken and what the conclusions are.

Matters arising are brief comments (up to 500 words) on articles and letters recently published in ***Nature***. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted.

Manuscripts may be submitted either to London or Washington. Manuscripts should be typed (double spacing) on one side of the paper only. Four copies are required, each accompanied by copies of lettered artwork. No title should exceed 80 characters in length. Reference lists, figure legends, etc, should be on separate sheets, all of which should be numbered. Abbreviations, symbols, units, etc, should be identified on one copy of the manuscript at their first appearance.

References should appear sequentially indicated by superscripts, in the text and should be abbreviated according to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963-65). The first and last page numbers of each reference should be cited. Reference to books should clearly indicate the publisher and the date and place of publication. Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text.

Each piece of artwork should be clearly marked with the author's name and the figure number. Original artwork should be unlettered. Suggestions for cover illustrations are welcome. Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published.

Important: Manuscripts or proofs sent by air courier should be declared as "manuscripts" and "value $5" to prevent the imposition of Customs duty and Value Added Tax in the United Kingdom.

GENECLEAN™

WHY SHOULD IT BE SO TIME-CONSUMING OR DIFFICULT OR EXPENSIVE TO DESALT OR PURIFY DNA SOLUTIONS, OR REMOVE DNA FROM AGAROSE?

IT SHOULDN'T . . .

AND NOW IT ISN'T. BIO 101 INTRODUCES **GENECLEAN™** A "QUICKIT" THAT ENABLES THE MOLECULAR BIOLOGIST TO:

- ☑ **REMOVE AND PURIFY DNA FROM AGAROSE GELS FAST** ... 20 minutes from Gel-Band to purified DNA in water or TE.
- ☑ **REMOVE UNREACTED TRACER-NUCLEOTIDES FROM LABELING REACTIONS and/or DESALT AND CONCENTRATE DNA SOLUTIONS, EVEN THOSE IN PHOSPHATE, FAST** ... 15 minutes without columns, alcohol precipitations or organic extractions.
- ☑ **ELIMINATE ENZYME-INHIBITING IMPURITIES FROM DNA SOLUTIONS FAST** ... 15 minutes to remove protein, small RNA and other enzyme-inhibiting impurities from DNA solutions. For example, remove RNA and protein from crude mini-plasmid cleared-lysates without protease, RNase, organic extractions or alcohol precipitations.
- ☑ **ELIMINATE NEED FOR ALCOHOL PRECIPITATION OF DNA** ... 15 minute **GENECLEAN** process serves the same purpose as alcohol precipitation of DNA, often with increased recovery.
- ☑ **REMOVE TRACES OF ORGANIC SOLVENTS FAST** ... 15 minutes to remove residual enzyme-inhibiting phenol, chloroform or ether from extracted solutions of DNA.
- ☑ **"QUICKCLONE" FROM HIGH MELTING-POINT AGAROSE GELS.**

BIO 101

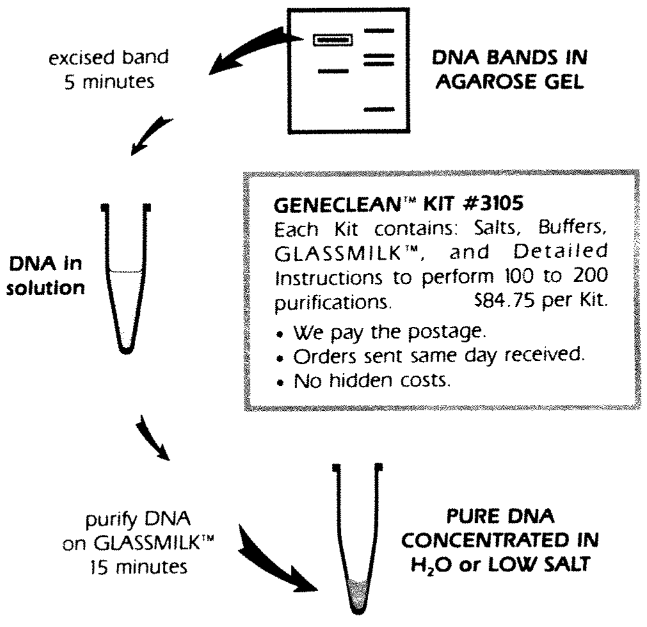
"DURING THE COURSE OF DOING RESEARCH, CHECK WITH BIO 101 ... WE HAVE YOUR TIME IN MIND."

MAILING ADDRESS:
BIO 101, INC., P.O. BOX 2284, LA JOLLA, CALIFORNIA 92038-2284

LOCATION:
4206-H SORRENTO VALLEY BLVD., SAN DIEGO, CA 92121

DNA TREATED WITH GENECLEAN IS A READY SUBSTRATE FOR MODIFYING ENZYMES ... without further manipulation.

GENECLEAN™ provides excellent recovery of single and double stranded DNA and it is not inhibited by SDS, organic solvents or protein.



excised band
5 minutes
DNA BANDS IN AGAROSE GEL
DNA in solution
purify DNA on GLASSMILK™ 15 minutes
PURE DNA CONCENTRATED IN H₂O or LOW SALT

GENECLEAN™ KIT #3105
Each Kit contains: Salts, Buffers, GLASSMILK™, and Detailed Instructions to perform 100 to 200 purifications. $84.75 per Kit.

- We pay the postage.
- Orders sent same day received.
- No hidden costs.

BIO 101 WANTS YOU TO BE PLEASED WITH ITS PRODUCTS. If you are not satisfied with the performance of **GENECLEAN™** return it within 30 days and we will refund your payment immediately.

To order **GENECLEAN™** Kit or for further information:

TOLL FREE Outside California, 1-800-424-6101
COLLECT from California, Canada 1-619-546-0556
TELEX 4993571 WORDS UI

Introductory discount on first orders; up to 15% on bulk orders!

Reader Service No. 12

nature

NATURE VOL. 324 18/25 DECEMBER 1986

nature
INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
CN JETS FROM COMET HALLEY

OPINION

How to give good dogs bad names 601
Slogans for recovery 602
Solstitial considerations

NEWS

Nuclear reprocessing 603
Chernobyl
European communities 604
Embryo research
Medical research 605
Japanese volcanoes
Pesticides 606
UK industrial research
Nuclear safety 607
Albanian outlook
Australia 608
Britain gets Cray supercomputer at last...
Rabies vaccine 609
Israel

CORRESPONDENCE

Wistar's export to Argentina/ Defensive deterrent/ South Africa 610

NEWS AND VIEWS

AIDS in an African context
Peter Newmark 611

Fuel technology: Control of nitrogen oxides
A Williams 612

A second messenger function for inositol tetrakisphosphate
Bob Michell 613

Quantum fluctuations in sub-micron wires
Patrick A Lee 613

A superfamily of potentially oncogenic hormone receptors
Stephen Green & Pierre Chambon 615

Views with a gravitational lens
Rachel Webster & Michael Fitchett 617

New data from mammalian homoeobox-containing genes
Michael H L Snow 618

Mitochondrial DNA: Rearranging the plant genome
Ronald A Butow 620

SCIENTIFIC CORRESPONDENCE

Microtubule dynamics
M W Kirschner & T Mitchison 621

Preservation in bogs
G R Dickson, E M Jope & A Boyde 622

How many reactor accidents?
J Schwartz 622

Alternative solutions
A R Brody, L H Overby & G M Janković 622

Archaeopteryx; the primordial bird
N Wickramasinghe & F Hoyle 622

BOOK REVIEWS

The Science Critic: A Critical Analysis of the Popular Presentation of Science *by M Goldsmith*
Frederick Dainton 623

The Induction of Hypnosis *by W E Edmonston, Jr*
Anthony W Clare 624

Marie Curie: A Life *by F Giroud*
H W Paul 624

Physiological Ecology of Animals: An Evolutionary Approach *by R M Sibly and P Calow*
Brian Charlesworth 625

Walkers *by M Jebb*
Eric Ashby 626

Biophilosophy: Analytic and Holistic Perspectives *by R Sattler*
P T Saunders 626

COMMENTARY

Forty years of genetic recombination in bacteria
A fortieth anniversary reminiscence
J Lederberg 627

Postmature scientific discovery?
H Zuckermann & J Lederberg 629

ARTICLES

Helium isotopes in sedimentary basins
E R Oxburgh, R K O'Nions & R I Hill 632

The c-*erb-A* protein is a high-affinity receptor for thyroid hormone
J Sap, A Muñoz, K Damm, Y Goldberg, J Ghysdael, A Leutz, H Beug & B Vennström 635

The c-*erb-A* gene encodes a thyroid hormone receptor
C Weinberger, C C Thompson, E S Ong, R Lebo, D J Gruol & R M Evans 641

LETTERS TO NATURE

Rotational period of comet Halley
R L Millis & D G Schleicher 646

Cyanogen jets in comet Halley
M F A'Hearn, S Hoban, P V Birch, C Bowers, R Martin, & D A Klinglesmith III 649

Nitric acid cloud formation in the cold Antarctic statosphere: a major cause for the springtime 'ozone hole'
P J Crutzen & F Arnold 651

Are domains in liquid crystalline polymers arrays of disclinations?
B A Wood & E L Thomas 655

Rapid reduction of nitrogen oxides in exhaust gas streams
R A Perry & D L Siebers 657

A magmatic heat pump
S A Morse 658

Formaldehyde oxime leaching of metals from deep-sea manganese nodules and other ores
R Derie 660

Homoeobox gene expression in mouse embryos varies with position by the primative streak stage
S J Gaunt, J R Miller, D J Powell & D Duboule 662

Differential and stage-related expression in embryonic tissues of a new human homoeobox gene
F Mavilio, A Simeone, A Giampaolo, A Faiella, V Zappavigna, D Acampora, G Poiana, G Russo, C Peschle & E Boncinelli 664

Contents continued overleaf

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Macmillan Journals Ltd and includes the Annual Index (mailed in February). Annual subscription for USA and Canada US $250. USA and Canadian orders to: *Nature*, Subscription Dept, PO Box 1501, Neptune, NJ 07753, USA. Other orders to *Nature*, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Second class postage paid at New York, NY 10012 and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by *Nature* to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of $1.00 a copy plus $0.10 a page is paid directly to CCC, 21 Congress Street, Salem, Massachusetts 01970, USA. Identification code for *Nature*: 0028-0836/86 $1.00 + $0.10. US Postmaster send address changes to: *Nature*, 65 Bleecker Street, New York, NY 10012. Published in Japan by Macmillan Shuppan K.K., Eikow Building, 1-10-9 Hongo, Bunkyo-ku, Tokyo, Japan. © 1986 Macmillan Journals Ltd.

Your **Source** *for* GROWTH FACTORS

pTGFβ_2

bFGF(b)

hPDGF

hTGFβ

ANTI TGFβ
(IgM) MONOCLONAL
(IgG) AFFINITY
PURIFIED

bFGF(a)

pPDGF

pTGFβ_1

I^{125} LABELED
GROWTH FACTORS

R&D
Systems

*T*he cytokine laboratories at *R and D Systems* isolate rare growth factors to the highest levels of purity and activity, and provide them at more affordable prices. All of our factors are assessed for purity by silver-stained gels, a.a. composition, and N-terminus sequencing. Activities are determined both in our own laboratories, as well as in those of established investigators in the field. To place an order, request product information or technical bulletins, call

CAT. NO.	FACTOR	PURITY	ACTIVITY (ED$_{50}$)
100-B	hTGFβ	>97	0.1–1.0 ng/ml[1]
101-B1	pTGFβ1	>97	0.1–1.0 ng/ml[1]
101-B2	pTGFβ2	>96	0.1–1.0 ng/ml[2]
120-HD	hPDGF	>95	1–2 ng/ml[3]
125-PD	pPDGF	>94	4–8 ng/ml[4]
132-FA	bFGF(a)	>96	1–4 ng/ml[5]
133-FB	bFGF(b)	>94	0.05–0.1 ng/ml[5]

1 (800) 328-2400

In Minnesota (612) 379-2956
Telex 750627

R&D

R & D Systems, Inc.
614 McKinley Place N.E.
Minneapolis, MN 55413

---

[1] Transforming growth factor type beta isolated from human platelets (hTGFβ) or from porcine platelets (pTGFβ1) is sequentially (first 50 a.a.) and biologically identical. Both give an ED$_{50}$ of 0.1-0.4 ng on NRK (49F) cells when assayed as per Assoian *et al* JBC *258*, 7155, 1983, and 0.8-1.0 ng/ml on AKR (2B) after Tucker *et al* CAN RES *43*, p1581, 1983.

[2] pTGFβ2 is a structurally distinct second form of TGFβ (sequence homology 70% in the first 50 a.a. to hTGFβ/pTGFβ1). When assayed as above, it gives the same ED$_{50}$s on NRK and AKR cells. However, some cell lines can distinguish between pTGFβ2 versus pTGFβ1/hTGFβ.

[3] Human platelet derived growth factor h(PDGF) is assayed on Balb/c-3T3 and NIH-3T3 cells for H^3-thymidine incorporation as per Raines and Ross, Meth. in Enz. *109*, p749.

[4] Porcine PDGF is structurally very different from hPDGF (i.e. differs in M.W., subunit composition, N-terminus), and is biologically less active in thymidine incorporation assays.

[5] Both the acidic and basic forms of bovine fibroblast growth factor are isolated from brain and assayed for H^3-thymidine incorporation on NR6-3T3 cells after Gospodarowicz, *et al* JBC, *253*, p3736, 1978.

Reader Service No.41

CONTENTS

SCIENCE IN IBERIA

Spain and Portugal have recently emerged from social and political upheavals: how has the practice of science been affected?

In this special issue (27 11 86) Nature publishes a survey by Alun Anderson of new developments in both countries, in which he comments on key issues affecting Iberian scientists

- ✻ What will Spain's new "law of science" achieve?
- ✻ Can Iberia reach the technological level needed to compete in the European Community?
- ✻ What have Iberia's researchers gained from their first decade of democracy?

Ordering information and prices as for back issues below

Annual Subscription Prices

UK		£104
Rest of Europe	Airspeed	£125
USA & Canada		US$250
Australia, NZ & S Africa	Airspeed	£150
	Airmail	£190
India	Airspeed	£120
	Airmail	£190
Japan	Airspeed	Y90000
Rest of World	Surface	£120
	Airmail	£180

Orders (with remittance) to

USA & Canada
Nature
Subscription Dept
PO Box 1501
Neptune
NJ 07753
USA

UK & Rest of World
Nature
Circulation Dept
Brunel Road
Basingstoke
Hants RG21 2XS, UK
Tel 0256 29242

(The addresses of Nature's editorial offices are shown facing the first editorial page)

Japanese subscription enquiries to
Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan
Tel (03) 292 3755

Personal subscription rates These are available in some countries to subscribers paying by personal cheque or credit card Details from

USA & Canada
Nature
65 Bleecker Street
New York
NY 10012, USA
Tel (212) 477-9600

UK & Europe
Felicity Parker
Nature
4 Little Essex Street
London WC2R 3LF, UK
Tel 01-836 6633

Back issues UK, £2 50, USA & Canada, US$6 00 (surface), US$9 00 (air), Rest of World, £3 00 (surface), £4 00 (air)

Binders
Single binders UK, £5 25 USA & Canada $11 50, Rest of World, £7 75 Set of four UK, £15 00, USA & Canada $32 00, Rest of World, £22 00

Annual indexes (1971-1985)
UK, £5 00 each, Rest of World, $10 00

Nature First Issue Facsimile
UK, £2 00, Rest of World (surface), $3 00, (air) $4 00

Nature in microform
Write to University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106, USA

Cross-regulatory interactions among the gap genes of *Drosophila*
H Jackle, D Tautz, R Schuh, E Seifert & R Lehmann 668

Location of a δ-subunit region determining ion transport through the acetylcholine channel
K Imoto, C Methfessel, B Sakmann, M Mishina, Y Mori, T Konno, K Fukuda, M, Kurasaki, H, Bujo, Y Fujita & S Numa 670

Abnormal tyrosine phosphorylation on T-cell receptor in lymphoproliferative disorders
L E Samelson, W F Davidson, H C Morse III & R D Klausner 674

Polymorphism of human Ia antigens generated by reciprocal intergenic exchange between two DR β loci
S Wu, T L Saunders & F H Bach 676

Predominant use of a V_α gene segment in mouse T-cell receptors for cytochrome *c*
A Winoto, J L Urban, N C Lan, J Goverman, L Hood & D Hansburg 679

High frequency of unequal recombination in pseudoautosomal region shown by proviral insertion in transgenic mouse
K Harbers, P Soriano, U Muller & R Jaenisch 682

***In vivo* protein–DNA interactions in a glucocorticoid response element require the presence of the hormone**
P B Becker, B Gloss, W Schmid, U Strahle & G Schutz 686

Steroid-free glucocorticoid receptor binds specifically to mouse mammary tumour virus DNA
T Willmann & M Beato 688

Molecular cloning and polymorphism of the human immune deficiency virus type 2
F Clavel, M Guyader, D Guétard, M Sallé, L Montagnier & M Alizon 691

A new way of enhancing the thermostability of proteases
T Imanaka, M Shibazaki & M Takagi 695

EXAFS study of the zinc-binding sites in the protein transcription factor IIIA
G P Diakun, L Fairall & A Klug 698

Site-directed mutagenesis reveals role of mobile arginine residue in lactate dehydrogenase catalysis
A R Clarke, D B Wigley, W N Chia, D Barstow, T Atkinson & J J Holbrook 699

Corrigenda . .

Protein kinase C activation induces conductance changes in *Hermissenda* photoreceptors like those seen in associative learning
J Farley & S Auerbach 702

A complimentary DNA clone for a macrophage-lymphocyte Fc receptor
V A Lewis, T Koch, H Plunter & I Mellman 702

Isolation of cDNA clones encoding the 20K non-glycosylated polypeptide chain of the human T-cell receptor/T3 complex
O P Gold, J M Puck, C L Pettey, M Cho, J Coligan, J N Woody & C Terhorst 702

NATURE CLASSIFIED

Professional appointments — Research posts — Studentships — Fellowships — Conferences — Courses — Seminars — Symposia: Back Pages

1 January in *Nature*:

- Optical digital computer
- Mitochondrial Eve
- Metal catalysts
- Vitamin D deficiency
- Chemical signalling in *Obelia*

GUIDE TO AUTHORS

Authors should be aware of the diversity of *Nature* s readership and should strive to be as widely understood as possible

Review articles should be accessible to the whole readership Most are commissioned, but unsolicited reviews are welcome (in which case prior consultation with the office is desirable)

Scientific articles are research reports whose conclusions are of general interest or which represent substantial advances of understanding The text should not exceed 3 000 words and six displayed items (figures plus tables) The article should include an italic heading of about 50 words

Letters to *Nature* are ordinarily 1,000 words long with no more than four displayed items The first paragraph (not exceeding 150 words) should say what the letter is about why the study it reports was undertaken and what the conclusions are

Matters arising are brief comments (up to 500 words) on articles and letters recently published in *Nature* The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted

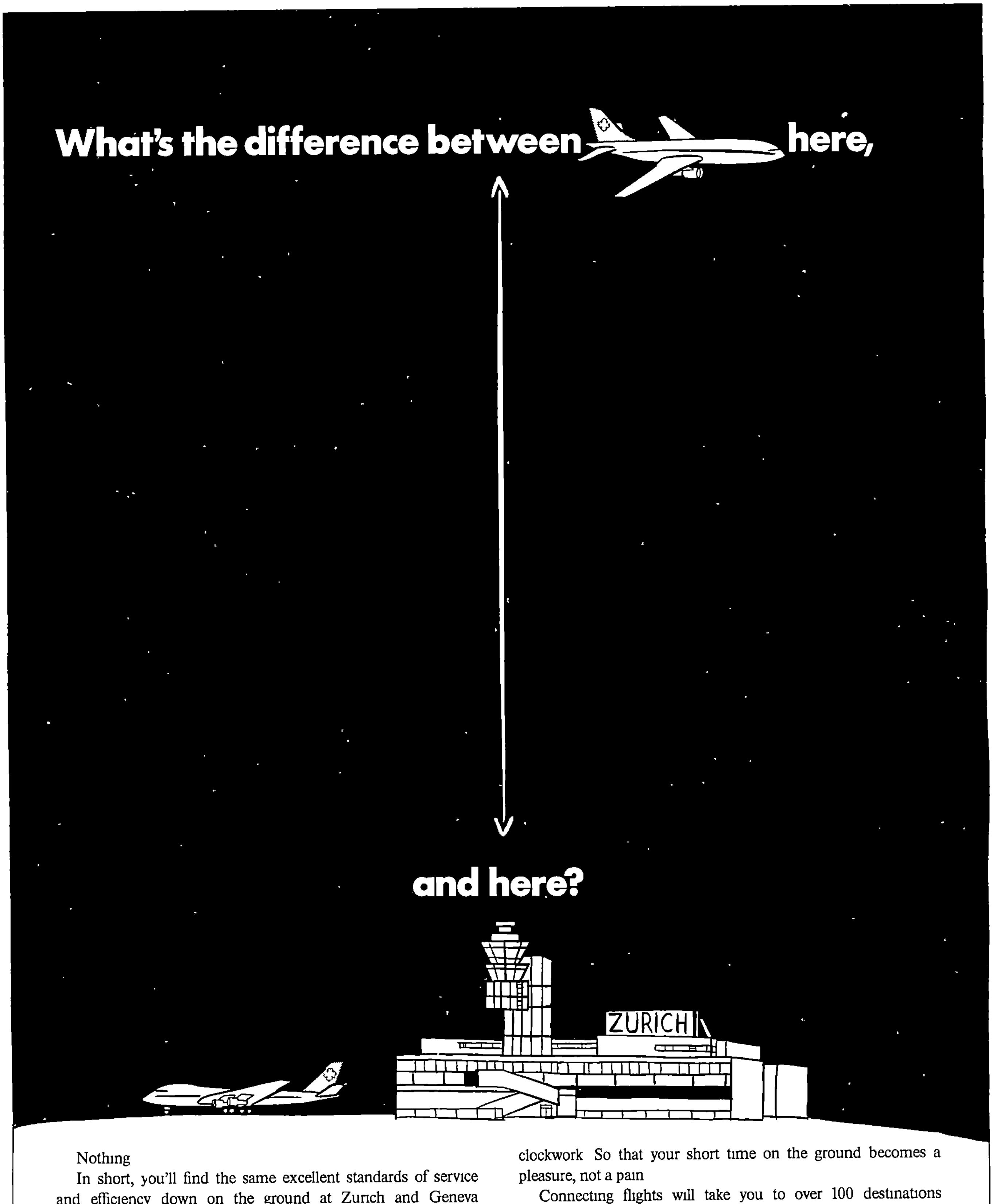
Manuscripts may be submitted either to London or Washington Manuscripts should be typed (double spacing) on one side of the paper only Four copies are required, each accompanied by copies of lettered artwork No title should exceed 80 characters in length Reference lists, figure legends, etc, should be on separate sheets, all of which should be numbered Abbreviations, symbols, units, etc, should be identified on one copy of the manuscript at their first appearance

References should appear sequentially indicated by superscripts, in the text and should be abbreviated according to the *World List of Scientific Periodicals*, fourth edition (Butterworth 1963 65) The first and last page numbers of each reference should be cited Reference to books should clearly indicate the publisher and the date and place of publication Unpublished articles should not be formally referred to unless accepted or submitted for publication, but may be mentioned in the text

Each piece of artwork should be clearly marked with the author s name and the figure number Original artwork should be unlettered Suggestions for cover illustrations are welcome Original artwork (and one copy of the manuscript) will be returned when manuscripts cannot be published

Important Manuscripts of proofs sent by air courier should be declared as manuscripts' and value $5 to prevent the imposition of Customs duty and Value Added Tax in the United Kingdom

Requests for permission to reproduce material from Nature *should be accompanied by a self addressed (and in the case of the United Kingdom and United States stamped) envelope*


What's the difference between here,
and here?
ZURICH

Nothing

In short, you'll find the same excellent standards of service and efficiency down on the ground at Zurich and Geneva airports as you're used to on a Swissair flight at 39,000 feet

There are always times when you can't fly direct to your long-haul destination and a transit stop can be a daunting prospect

At Zurich and Geneva you'll experience first rate amenities and an organisation that ensures that everything runs like clockwork So that your short time on the ground becomes a pleasure, not a pain

Connecting flights will take you to over 100 destinations worldwide, on which you can experience the level of service that's won us Business Traveller Magazine's Best Airline award every year since it started in 1980

Fly Swissair We make all the right connections

swissair

Reader Service No 58

nature

NATURE VOL 324 6 NOVEMBER 1986

What must be done about AIDS

Governments can no longer hope that AIDS will simply go away Time could yet show that AIDS is as great a threat as nuclear war. And the time has come to act But how?

THERE are two extreme views about the occurrence of AIDS (acquired immune deficiency syndrome) in places such as North America and Western Europe One, exemplified by the sponsors of Proposition 64 on this week's Californian ballot, is that AIDS has arisen because some social groups follow unusual sexual practices (homosexuality) or inject themselves with drugs, that AIDS would go away if these practices were stopped and that society should meanwhile protect itself by casting out the sick The other view is that AIDS has spread so rapidly since it was first recognized as a distinct disease in mid-1981, and is now spreading so rapidly outside the groups originally infected, that it can be only a matter of time before a substantial proportion of the general population, women and children as well as men, succumbs If the second view should prove correct, and if no generally applicable remedy or prophylactic comes on the scene, the damage done to modern society by AIDS would compare with that expected from, say, a nuclear war between the major powers

That is the background to the publication last week of the report of a committee sponsored jointly by the US National Academy of Sciences and the Institute of Medicine One of the two committee chairmen is Dr David Baltimore, director of the Whitehead Institute at Massachusetts Institute of Technology and, piquantly, one of the two independent discoverers (with Dr Howard Temin of the University of Wisconsin) of the role of the enzyme reverse transcriptase in the replication of viruses whose genetic information is carried by RNA, of which the AIDS virus is one The document (see p 3) is a model of clarity and careful assessment of daunting problems It denies the first and simplistic view of AIDS and how to deal with it Properly, it is agnostic on the second view of the distant future But that does not prevent it from urging, irresistibly it would seem, that the US government should vastly increase its spending on the AIDS epidemic In round numbers, it wants to see an extra $1,000 million a year spent on education and public health and as much again on research

The seriousness of the problem cannot be dodged The US Public Health Service estimates that 9,000 US citizens will have died of AIDS during 1986, doubling the cumulative toll at the beginning of the year During 1991, on the same set of estimates, there will be some 54,000 deaths in the United States, almost all of them recruited from the 1 to 1 5 million people thought already to have been infected Elsewhere, the pattern is much the same, but lags behind that in the United States by a year or two In Britain, for example, the number of known cases has doubled each year since 1983, and now amounts to just over 500, but there are some tens of thousands of infected people to fill the ranks of the dying in the years ahead

Plague

So is this the plague returned? Much depends on what is meant by that concept, of which one essential ingredient must be that there is little that can be done about it What should chill the spines of the US Congress elected this week is the judgement of the academies' committee that there is no quick therapeutic or prophylactic fix in sight, last month's wonder drug, AZT (azidothymidine) is at best a palliative that that could prove unacceptably toxic, while a vaccine against the virus responsible (now called HIV) is at least five years away Meanwhile, on the estimates of the US Public Health Service, AIDS will be responsible for between 1 and 2 per cent of deaths in the United States five years from now The Baltimore report is right to emphasize the great social and economic cost of that toll, with its peculiar age-distribution between early adulthood and middle age, but what is the chance that the death-rate might grow to be an order of magnitude greater? That is what the US and other governments will want to know While AIDS is so new, it is too soon to tell, but an answer will emerge over the next few years Two technical questions are crucial

Bad news

- What proportion of those asymptomatically infected will eventually succumb? Since the introduction two years ago of blood tests for the detection of antibodies against HIV, it has become clear that infection is not immediately followed by AIDS This is most directly shown by the retrospective examination of blood collected both in California and Central Africa in 1978, three years before AIDS was recognized for what it is, yet showing evidence of antibodies The US Public Health Service's estimate of the development of AIDS in the United States supposes that 20–30 per cent of those with antibodies will convert to overt AIDS during the next five years, but that is but an educated guess based on five years experience of AIDS, only two of them with the benefit of antibody tests The academies' committee prefers the estimate of 25–50 per cent of conversion to overt AIDS within 5–10 years of the appearance of antibodies in the blood But what if the conversion rate is more like 100 per cent over, say, two decades? Nothing in the data so far collected can exclude that possibility, which would be very bad news
- How effective is heterosexual intercourse as a means of infection? There was a sprinkling of women among the first few handfuls of AIDS cases reported in the United States, but these could at the outset be explained as exceptions to the rule that AIDS is a venereal disease of male homosexuals (or else is acquired by blood transfusion, haemophilia treatment or the use of second-hand syringes for injecting illicit drugs) But this view, never plausible, is no longer tenable The academies' committee refers to substantiated cases in the United States of heterosexual transmission from men to women and in the other direction In many African countries, as well as in Haiti in the Caribbean, women and men are more equally infected In due course, if the social groups now considered to be especially at risk (male homosexuals, intravenous drug users and users of blood products) are set to one side, that will be the case in advanced societies as well What will matter is the degree to which the infection spreads among the heterosexual population free from other complications The academies' committee says that at present there is no worthwhile evidence bearing on this question But if the answer should be gloomy, the implication could be that mere containment of the spread will be impractical That, too, would be exceedingly bad news

So how is it possible to fight an infection potentially so dangerous for which, now, there is neither prophylactic nor cure? The academies' committee agrees with the drafters of Proposition 64, which would make AIDS a notifiable disease, on just one point in the short term, the only possible stratagem is to reduce the rate at which the infection can spread (Historians will note the humiliating similarity with the stratagems used to contain the Black Death) Then two two groups part company

The committee explicitly rejects proposals for the compulsory testing of the social groups considered most at risk (followed, perhaps, by the compulsory incarceration of those found positive in quarantine camps) not only on libertarian grounds but, more cogently, on the grounds that such hare-brained schemes could not work

There are two grounds for this position First, there is a variable, undetermined and perhaps long interval after infection before the appearance of antibodies (seroconversion) to HIV in a person's blood That means that even ideally efficient tests could not identify all infectious carriers More important, there is ample evidence that attempts at coercion in this field would be counter-productive, even now, without it, the social stigma of being infected with HIV means that some patients fail to report to their physicians and, for understandable if unforgivable motives, that physicians fail to report cases in their care to the authorities Compulsion would be a way of making the disease seem to go away by driving it underground where it would be an even greater danger

The academies' committee instead puts its money (or, rather, the federal government's) on education, demanding urgent efforts to enlighten not merely the groups at risk but the other members of the communities to which AIDS has or will spread There is some good news General enlightenment has taken the edge off the fear, rife even two years ago, that a person could catch AIDS from being in the same room as a patient There is also some sketchy evidence that the behaviour of homosexual males has been changed by the spread of an understanding of the disease among those at risk, much of that accomplished by voluntary self-help groups (Even so, the committee cannot decide whether the recent flattening of the rate of infection among homosexuals in New York City and California is a consequence of changed habits or, equivalently, the saturation of the infective population) The committee is right to say that there are two other groups to which public education should be directed intravenous drug users (who seem mostly beyond the ken of most social institutions) and those still at school on the threshold of adult sexual life The committee might have made more of the threat posed by prostitutes

Irony

These are daunting problems, loaded with irony If the new US Congress takes the academies' report seriously, as it must, it will find itself having to spend as much on AIDS education as its predecessor chose to appropriate for the great war on drugs At least it may be some comfort that, in respect of the transfer of HIV by means of used hypodermic needles, the two wars will overlap Yet the committee is right to emphasize that, with the practice of screening blood products to eliminate blood from infected donors and the declining pace of new infection among male homosexuals, the intravenous drug users are the next most important reservoir of infection Who will be able to reach them?

The young, even though they turn up at schools most days, are in some ways even less approachable AIDS is a venereal disease that may be acquired in the course of sexual intercourse of any kind, but how are young people without sexual experience to be warned that it may be a dangerous as well as a natural practice? The committee comes out squarely for explicit sex education in the public schools and for the replacement of euphemisms such as 'intimate contact' by frank statements of what their users mean, in the vernacular if need be That will mean more classroom four-letter words

The committee has no choice, yet this is where its troubles will begin, it acknowledges that "political opposition and bureaucratic intransigence" at present limited school sex education to a few enlightened school jurisdictions, but seems not to appreciate that the opposition is as likely to harden as to melt away, now that there is AIDS as well as all those other problems British readers of the report will know the difficulties, for this week in which their government will have placated its supporters opposed to sex education by transferring from local education authorities and head teachers to supposedly less free-thinking school governors the right to decide on sex education in the curriculum Fair play, the politicians are in a nasty cleft stick

The stakes are high, they risk that AIDS might spread beyond the point at which it could be contained Their constituents, as always ready to believe that it would be simpler to get from A (now) to B (tomorrow) by starting from somewhere else, say they would prefer that teachers on the taxpayers' payroll should tell innocent children about the virtues of the nuclear family than about the condom as a barrier to venereal infection But they, too, have no choice but to fall in with the logic of the academies' conclusion that only the most explicit sex education for the young offers hope of avoiding what may be a catastrophe Issuing chastity belts for all children would be the other way

From all this, it is possible that some good news may emerge The remarkable feature of AIDS infection in the past five years is that so much has been learned about it so quickly It is, for example, great good luck that the discovery of reverse transcriptase preceded the emergence of AIDS as a disease It is also remarkable that so much has been done, in a mere three years, to identify and characterize the virus The blood test, with all its defects, helps Yet the provision of a prophylactic, let alone of a cure, will require a more sensitive and complete understanding of the human immune system than there is at present The committee makes a strong case for giving research a chance, will governments agree?

Ignorance

Maybe, but maybe not In Washington, the congressional committees will call the shots, they, not the likes of Baltimore, have the cheque-books in their fists But congressmen will not be pleased, early in 1987, to be confronted with the demand of the committee academies' $2,000 million a year for the sake of saving civilization, especially when the Strategic Defense Initiative will be asking for even more, backed by a tighter prospectus, for ostensibly the same purpose The best way to choose between projects such as these is to calculate the downside risks If SDI should come to nothing, the world would probably be no more dangerous than it is (and might even be safer), but if the search for treatment and prophylaxis for AIDS should fail, calamity will be on the cards

The next few years, while there accumulate the data on which it will be possible to estimate the eventual penetration of the general population, will be crucial It is a brief breathing space, one in which intelligent policies could help decelerate the spread of the disease and pave the way for techniques that will be needed, in any case, for decades to come It will be a curious period for us all, not only for politicians Somehow it will be necessary to combine a sense of urgency with a sense of the importance of the civility that makes society worthwhile The danger is that some will be overwhelmed by the urgency of the problem to the point at which illiberal ways take hold The latent and not-so-latent hostility to homosexuals is a constant danger So are groundless attempts to keep children with AIDS out of the public schools Heavy hints by insurance companies (as in Britain at the weekend) that voluntary blood tests will be taken as presumptive evidence of sexual risk, are both inequitable and counter-productive What we must remember is that, as in the defence of Western Europe, battles that destroy what they are intended to preserve are not worth winning □

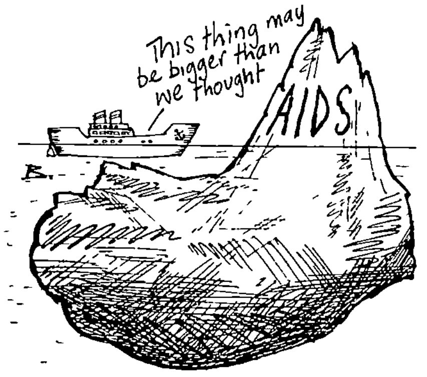
AIDS

Academies demand urgent public education

Washington

AN immediate and massive public education campaign is needed to slow the spread of acquired immune deficiency syndrome (AIDS), according to a report released last week by the Institute of Medicine (IOM) and the National Academy of Sciences* In addition to the $1,000 million a year required to pay for this educational campaign, the report also urges that support for AIDS research should be increased to $1,000 million a year by 1990 Without decisive action, IOM president Samuel Thier says, the epidemic may "snowball into an even greater catastrophe"

Following it annual meeting in 1985, the Institute of Medicine determined that it should provide direction for the US effort to combat AIDS With Thier as its new president, the institute together with the academy made the development of recommendations for dealing with the AIDS epidemic a priority Last week's 374-page document was produced in slightly more than seven months David Baltimore of the Whitehead Institute and Sheldon Wolff of Tufts University Medical School co-chaired the panels responsible


This thing may be bigger than we thought
AIDS

Although the report praises government support for research on the biology of the AIDS virus (HIV, or human immunodeficiency virus), the report is critical of the government's effort in public health In particular, federal educational efforts are dubbed "woefully inadequate" Until an effective vaccine is developed, Wolff says that educating the public is the best hope for curbing the disease The report reckons an effective campaign involving all levels of government as well as industry and private organizations will cost about $1,000 million a year

Almost as if anticipating the IOM report's criticism, Surgeon General C Everett Koop called for an increased educational effort in a report released one week before the IOM report Declaring that "the silence must end", Koop says it is no longer possible "to sidestep frank, open discussions about sexual practices — homosexual and heterosexual" Education must be extended to prepubescent children as well as to their parents, according to the Koop report

Wolff describes the surgeon general's report as "clearly a step in the right direction" Both reports urge the use of condoms for men who are not certain if their partners are infected with HIV

A vexing problem in determining how best to orchestrate an educational campaign against the spread of AIDS is the lack of information about sexual practices Baltimore suggests that a significant portion of new research money should go to social science research in that area Sexual intercourse remains the most common means of transmitting HIV Although the disease primarily afflicts homosexuals, the Public Health Service estimates that heterosexual transmission will account for 10 per cent of new cases by 1991

As well as increased education, the report makes several specific recommendations about new public health efforts that should be attempted As non-sexual transmission of AIDS occurs primarily through sharing of needles used for injection of intravenous drugs, the government should investigate the possibility of providing clean needles free of charge to users of illegal drugs In strictly financial terms, the report says it is cheaper to treat a drug addict than an AIDS patient

The report also calls for increased surveillance for both AIDS and AIDS-related complex (ARC) cases, and voluntary testing of serological status for individuals who may have been exposed to the virus The report declares that mandatory blood tests are inappropriate in most cases, the armed forces being a notable exception Calling AIDS an international problem, the IOM report urges the United States to spend $50 million a year on efforts to control the disease throughout the world

The report recommends that a new National AIDS Commission be established to coordinate federal efforts in combating AIDS But Baltimore does not see this office as playing a role in directing research He says that research proposals initiated by investigators make the best use of scientific talent in learning more about HIV, but that a directed approach would be appropriate for a vaccine development project

Federal funding for AIDS research has been nearly doubling in recent years In 1984, the government spent just over $61 million on AIDS research For 1987, that figure will jump to $410 million Gary Noble, now coordinating AIDS activities for the Public Health Service, will not speculate whether that rate of increase will be sustained Noble agrees in principle with many of the IOM report's conclusions, but says new therapeutics or unexpected discoveries about HIV could change matters

By 1991 the Public Health Service estimates there will be 270,000 cases of AIDS in the United States There are already an estimated 1 5 million people infected with HIV Academy president Frank Press says there is a need for a "national strategy to respond to a national crisis" Thier hopes the reports will mark the "turning point in the nation's approach to AIDS"

Joseph Palca

---

* *Confronting AIDS Directions for public health, health care and research* National Academy Press, Washington, DC 1986

AIDS drug test

SHROUDED by a veil of patent secrecy, an antiviral agent that is claimed to have more laboratory activity against the AIDS (acquired immune deficiency syndrome) virus than other anti virals is about to undergo animal toxicity tests in Britain. Porton International, the company that is developing the drug, hopes that it will pass the toxicity tests and can rapidly be tested by patients, using the US Food and Drug Administration's special fast-track procedure for potential AIDS drugs.

Porton will not, however, reveal anything of the chemical nature of the drug, which is named human immune virus antiviral. It is a natural substance extracted from a microorganism, and first emerged during a screening for antiviral compounds at the University of California (exactly where and by whom, Porton will also not reveal). Porton has acquired the rights to develop the drug which it is doing in collaboration with the Centre for Applied Microbiology and Research at Porton Down in Wiltshire, known for its expertise in the large-scale culture of microorganisms

Peter Newmark

- THE British government, which this week set up a cabinet committee under Lord Whitelaw to look into the problems posed by AIDS is also under pressure to increase the publicity given to it and to consider a mail-drop to every home in the United Kingdom

Critics of government strategy claim that the publicity campaign conducted in the British press since the spring has been a failure, as a high level of ignorance about the disease still prevails Advertisements were placed in national newspapers this year, and AIDS publicity has been funded through a £2 5 million budget ☐

Test-ban seismology

Soviet visit to US cut short

Washington

THE unique private effort to monitor nuclear testing in the United States and the Soviet Union is off and running, with three seismic monitoring stations now established near the Soviet test facility in Semipalatinsk But phase two, due to start next week with the arrival of five Soviet seismologists for site selection in this country, will proceed as planned The US State Department has limited the Soviet delegation's stay to one week, and has refused to permit a planned visit to potential monitoring sites

The monitoring effort comes from an agreement signed last spring by the Natural Resources Defense Council (NRDC), a private non-profit organization, and the Soviet Academy of Sciences (see *Nature* **321**, 638, 1986) Phase one of the agreement was completed this summer when NRDC scientists established seismic monitoring stations at Karkaralinsk, Bayanaul and Karasu near Semipalatinsk

Preliminary measurements have been taken at each site, and construction is nearly completed on 300-ft boreholes for high-frequency seismometers Returning last week from a visit to the Soviet sites, NRDC senior scientist Thomas Cochran declared the Soviets had done "an exceptional job, spending some $150,000 for construction of each station"

The US government, by contrast, has not welcomed the NRDC effort The administration feels that arms control issues should be carried out on a government-to-government basis, not by private organizations The goal of the NRDC/Soviet Academy effort is "to get the verification issue out the way", says Cochran But the State Department believes that there is a better way to monitor nuclear testing The government's approach is a a hydrodynamic yield measurement technique called CORRTEX, for Continuous Reflectometry for Radius versus Time Experiment

A coaxial cable is placed in a hole parallel to the nuclear device After the explosion, the rate of change in the cable length is used to estimate the yield of the device President Reagan last March offered to demonstrate CORRTEX for Soviet scientists, but so far that offer has not been accepted The State Department made viewing a CORRTEX demonstration a condition for granting the two-week visa sought by the Soviet seismologists, but Cochran says that Soviet officials decided that the seismologists would be "inappropriate" for determining the value of the system

During their visit, the Soviet delegation will also visit Teledyne Geotech and Refraction Technology in Dallas, manufacturers of monitoring equipment for the monitoring effort NRDC expects to take delivery of the new equipment later this month, and after testing it should be installed on site in the Soviet Union by early next year The Soviet seismologists will also discuss with seismologists at the Scripps Institution at the University of California, San Diego, the selection of monitoring sites near the US Nevada test facility Three potential sites have been identified at Nelson and Railroad Valley, both in Nevada, and Deep Springs, California

Joseph Palca

British seek funds

THE British team, invited by the Soviets to join the Americans in monitoring nuclear tests at Semipalatinsk, is attempting quickly to raise the bulk of the £200,000 it will need for the coming year.

Two committees will steer the project, one composed of eminent scientists with a reputation for raising money and the other of seismologists Seismologists from the Universities of Leeds, Liverpool and Leicester have already offered to help.

The invitation was extended by the Soviet Academy early last month to two British scientists visiting Moscow, Dr Frank Barnaby, former director of the Stockholm International Peace Research Institute, and Dr Jeremy Leggatt, lecturer in Earth sciences at Imperial College in London.

Apart from Barnaby and Leggatt, the project team includes Dr David Davies, a seismologist and former editor of *Nature* who is now director of the Dartington North Devon Trust, Professor Martin Rees of the Institute of Astronomy, Cambridge, and Joseph Rotblat, emeritus professor of physics at St Bartholomew's Hospital Medical School in London.

The British team should be in place by mid-December, before the test-ban treaty lapses at the end of the year. If the group can attract sufficient funds, two monitoring positions will be created, manned by two people at any one time.

Because of the political aspects of the project, no research council support will be available, Leggatt says, in spite of the scientific value of the proposed collaboration between British and Soviet seismologists. It is hoped, however, that Quaker and US foundations will provide support.

The equipment will all be supplied from the West, although the Soviets will have access to data retrieved. The British monitoring positions will collect data on the crustal struture below the test sites, particularly seismic wave propagation

Bill Johnstone

Strategic Defense Initiative

Academicians doubt efficacy

Washington

SCIENTISTS at Cornell University announced last week the results of a commissioned survey on views held by members of the National Academy of Sciences about the Strategic Defense Initiative (SDI) The principal conclusion is that 78 per cent of those responding believed that the prospects were "poor" or "extremely poor" that SDI can be made "survivable and cost effective at the margin" within the next 25 years, while only 4 per cent believed the chances of success were better than even Ninety-eight per cent believed SDI could not provide an effective defence of the US population within 25 years in the face of Soviet attempts to thwart it, and 94 per cent believed that goal to be impossible even if the Soviets froze their arsenals at current levels

The survey has caused some discomfort at the academy because it does not represent an official view, although the Cornell scientists have been careful to make the distinction clear, academy staff feel it is now perceived as opposed to SDI The study, designed and conducted by the Cornell Institute for Social and Economic Research, was instigated by Peter Stein, a physics professor at Cornell, and paid for by the Ploughshares Fund of San Francisco and the MacArthur Foundation The survey was sent to all 663 academy members in the physical, mathematical and engineering sciences, and responses were received from 71 per cent

The study nevertheless received some influential support from academy members A covering letter sent out with the questionnaire, signed by well-known Cornell members, urged recipients to give a few minutes to respond, saying that a "definitive poll" could provide a "useful contribution to one of the most vital science policy issues of our time"

The administration, well used to academics querying SDI, is unlikely to be immediately influenced by the survey But Congress is another matter it recently voted $3,500 million for SDI research in fiscal year 1987, $1,900 million less than the administration requested But for a one-vote close call, the figure would have been only $3,300 million, and doubts about SDI's feasibility played a significant part in reducing the budget Senator William Proxmire (Democrat, Wisconsin), a prominent SDI critic who was invited to the press conference at which the survey was announced, said "the best scientific minds of the country are telling [the President] loud and clear that Star Wars won't work and that it's a waste of money"

Tim Beardsley

Robots

Britain tries to catch up

BRITAIN is to have an Advanced Robotics Research Centre to develop new techniques for automatic manufacture and to cultivate the wider acceptance of robot technology, in which Britain still lags behind many of its major industrial competitors.

The announcement of this development coincides with the publication on 5 November of a study of the uses of robots in Britain, which concludes that there are still many fewer than 1,000 manufacturing companies using robots, roughly one robot for every 40 factories. The robot population of Britain at the beginning of the year is estimated at 3,200, fewer than the annual increase of robots in West Germany alone during 1985.

The new research centre will be supported jointly by the British government and between 20 and 25 industrial partners. Possible locations include the universities of Glasgow, Salford and Edinburgh and the UK Atomic Energy Authority research establishment at Harwell. The budget will be determined by the number of industrial partners in the venture.

The plan is that the institute will have a research staff of two dozen, concentrating on advanced manipulators, sensors (visual, tactile and acoustic), navigational systems for self-propelled devices, computer systems software and standards for control and artificial intelligence or Intelligent Knowledge Based systems. The new group will also investigate advanced designs intended for use in hostile environments unsuitable for people, for example, nuclear installations.

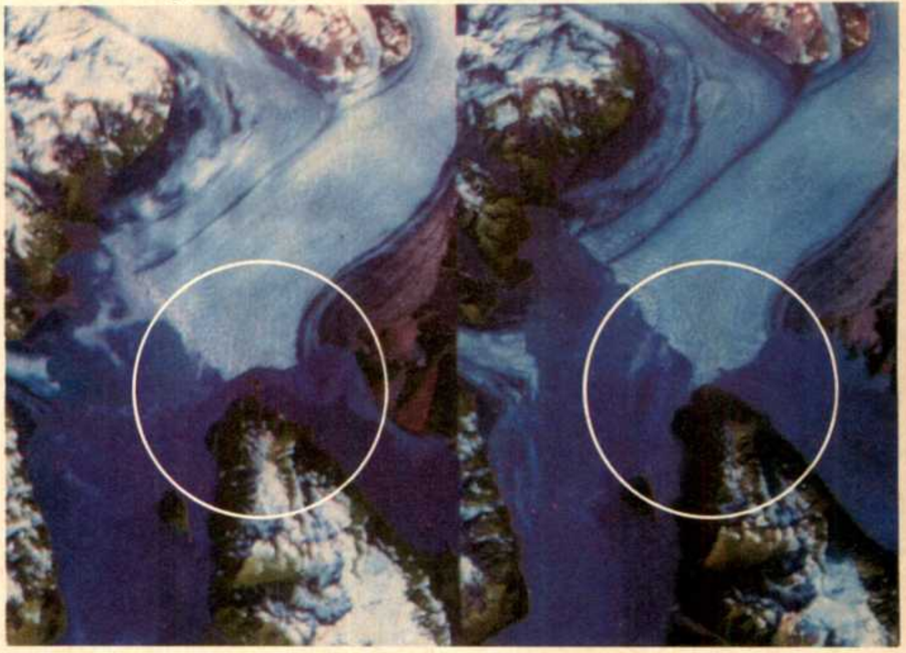
The study published last week, which advocates the setting up of a centre along the lines now proposed, was carried out by the London-based Policy Studies Institute on the initiative of the British Robot Association with support from the Gatsby Charitable Foundation, the Department of Trade and Industry and the National Economic Development Office.

The report says that there were industrial robots in about 740 manufacturing plants at the beginning of 1986, but that most of the plants belonged to larger companies. Only 12 per cent of robot users employ fewer than 100 people. On the other hand, a third of the users are very large manufacturing plants employing more than 1,000 people.

The report says that part of the problem in Britain has been the high cost of the development of robots, the lack of technical support and the poor reliability of designs. The new robot centre is meant to offer a cure for some of these ills.

Bill Johnstone

Alaska glacier seals fate of fiord wildlife



WHEN the Landsat 5 image on the left was taken on 7 August last year, water from the Russell Fiord in Alaska (on the right in each image) could flow freely into Disenchantment Bay (left). But by 11 September this year, as shown in the image on the right, the Hubbard Glacier had sealed off the fiord, forming Russell Lake. Movement of the glacier could eventually fill Disenchantment Bay and the larger Yakutat Bay with ice.

This event is "probably the largest natural alteration in ocean, glaciers, lake and rivers in North America within our lifetimes", says US Geological Survey glaciologist Larry Mayo. The inhabitants of Yakutat on the Pacific are probably not in any danger, but saltwater fish and wildlife, including seals, trapped in Russell Lake will almost certainly die as the water becomes deoxygenated.

In these images vegetation shows up geen, rock and oils a red and water/ice are in blue and white shades. □

US education

Plan to license teachers

Washington

THE US National Science Teachers Association (NSTA) is planning a certification scheme to establish basic qualifications for all science teachers. The plan has been prompted by a study showing that half of newly hired teachers in the United States are unqualified and that about a third of all science classes are staffed by unqualified teachers.

The immediate problem is that state certification requirements are now" largely ignored" in hiring and assignment decisions, according to Bill Aldridge, executive director of the association. Although participation in the certification scheme will be voluntary, Aldridge hopes that it will put pressure on states to improve and enforce their own certification standards.

The association also hopes that a further benefit of its scheme will be that specialist teachers in high schools will be more accurately assigned to suitable classes. Another survey by NSTA showed that 65 per cent of physics teachers, 52 per cent of chemistry teachers and 37 per cent of biology teachers are assigned to classes other than those they are qualified to teach.

To qualify for NSTA certification, teachers will have to hold an appropriate bachelor's degree and will have to demonstrate a strong background in statistics, mathematics and the application of computers to science teaching. In-service training and three years experience will also be required.

NSTA's standards have been adopted by the National Council for the Accreditation of Teacher Education, and will be the basis for future decisions on the accreditation of college and university science teacher training programmes.

School administrators are likely to welcome the certification scheme, although Gary Marks of the American Association of School Administrators is concerned that the pressure to raise standards may impose greater financial burdens on school systems. He also asks whether, if certification becomes standard, school administrators will be forced to drop courses when certified teachers are not readily available.

Carol Ezzell

US oceanography

Research centre born in storm

Washington

THE US Navy's new Institute for Naval Oceanography (INO) in Bay St Louis, Mississippi, opened last month with ambitions of becoming the leading ocean modelling research centre in the 1990s. But critics are still bitter at the decision by the Navy to put the institute at the unpopular Mississippi site rather than at Monterey, California, the choice of the great majority of the experts consulted. Some say that the Mississippi site, which is far from any major university and suffers from extreme heat and humidity in the summer, will be a "major impediment" to hiring the best researchers in the field.

According to consultants who worked on the project, Bay St Louis was not even the first choice of Secretary of the Navy John Lehman. The final decision was not made until 18 months ago, and many see it as reflecting the influence of Mississippi Senator John Stennis, a powerful Democrat on the Senate Armed Services Committee. Stennis is seen as the force behind a long-term agglomeration of Navy oceanographic facilities at Bay St Louis.

Some oceanographers say that opportunities for INO collaboration with the Navy's Fleet Numerical Oceanographic Center and its postgraduate school at Monterey would have been better than with the Navy Oceanographic Research and Development Activity (NORDA), already at the Bay St Louis site. (Researchers staged a protest several years ago when part of the Naval Research Laboratory in Washington was moved to Bay St Louis to become part of NORDA.) And the Monterey climate is as famed for being pleasant as that of the Mississippi delta is for being unpleasant.

Even so, a member of Stennis's staff denies that political pressure had been brought to bear, saying the decision reflected priorities decided by the Navy several years ago.

INO will work on all aspects of ocean modelling and prediction, using supercomputers at other centres until it acquires in about two years its own 'class 7' supercomputer, the generation beyond the most advanced Cray machine, the XMP 4800.

Although the centre was formed primarily to improve predictions for naval operations, its director, Christopher Mooers, foresees many spin-offs to civilian applications. At present, ocean predictions are limited by atmospheric variability, and a major goal will be to integrate atmospheric data with oceanic models. Mooers says that working at INO will "appeal to those with a pioneering spirit".

The immediate goal, for the year or so ahead, is to provide regional predictions. The centre will be fully operational in the early 1990s. The institute is timed to make best use of the expected welter of new remote-sensing ocean data in the next decade, when the National Aeronautics and Space Administration's TOPEX satellite will be in service and the US Navy launches its NROSS oceanographic satellites. Walter Munk, chairman of the National Academy of Sciences' ocean studies board, endorses the formation of the institute, saying of ocean modelling that "no-one knows yet how well it can be done". INO will also integrate data from acoustic sensors and *in situ* measurements.

INO has been established at the National Space Technology Laboratories, a remnant of the Apollo space programme, with an initial budget of $4 million and a staff of 30. It is expected to expand to about 50 within five years, and will set aside about $1 million each year to support work at other institutions. Mooers specifically welcomes proposals for collaborative research with overseas scientists.

Many academic oceanographers say quality suffers at Navy-run establishments, partly because military directors are often on short-term appointments. Mooers says that the NCAR consortium "provides our academic quality assurance". But other oceanographers are more cautious, saying the institute's potential to become a world leader will be met only if the Navy resists the temptation to keep it focused on short-term goals.

Tim Beardsley

Soviet Union

Russian criticism begins at home

THE election of Academician Gurii Marchuk as president of the Soviet Academy of Sciences has been celebrated with massive self-criticism. Instead of dwelling on recent achievements, both the report on the state of the academy presented by vice-president Vladimir Kotel'nikov and the ensuing discussion stressed deficiencies in performance ranging from overstaffing with inept personnel to a shortage of computers of "average productivity".

In some cases, although no names were mentioned, the faults seem to have been at least in part associated with the attitude of the outgoing president, Anatolii Aleksandrov, towards Western science. It was Aleksandrov's contention that Soviet science can achieve world leadership in all fields on its own efforts, and that Soviet participation in international cooperation programmes should be undertaken only to further detente and to avoid duplication of effort. Now, however, Kotel'nikov's report stresses that the CoCom restrictions on the transfer of Western state-of-the-art instruments constitutes a serious problem and the Soviet instrument manufacturing industry cannot satisfy the needs of the scientific community. Some time ago, said Kotel'nikov, the academy raised the question of how to expand the development and manufacture of scientific instruments but it has made little progress so far, although capital investment in the academy's own instrument-making base is to increase by 150 per cent during the present five-year plan.

Similarly, Academician Boris Paton, president of the Ukrainian Academy of Sciences, urged the importance for Soviet planners of a proper understanding of technological developments in the West. Although technology is "the most holy of holies to which there is virtually no access", Soviet scientists must use all "possible and, so to say, impossible" means of establishing this information.

Paton's suggestions raise a number of interesting questions. The Soviet Union already has an extensive information network, administered, in the civil sector, by the State Committee for Science and Technology, and in the military sector by the Commission for the Defence Industry. Information on Western science and (in particular) technology, obtained by debriefing scientists returning from visits abroad, is fed into these twin networks through their permanent representatives in the research institutes. The reverse engineering of Western prototypes coming into Soviet hands is likewise carried out under the auspices of these bodies. What Paton appears to imply, therefore, is that either the academy scientists have not been sufficiently zealous in their efforts to acquire useful technological information, or that the analysis carried out by the existing information services is inadequate.

As far as Soviet science is concerned, most criticisms proved to be aspects of the endemic Soviet theme: the failure to implement research results in manufacturing and agricultural practice.

But some practical suggestions were proposed. Academician Vladimir Tuchkevich noted that, although academic resesearch institutes are encouraged to serve the needs of industry, they have no financial incentive to do so. Instead, the research is financed by the state, and the results, including any hardware produced by the research team, either have to be handed over free of charge or sold at less than cost price. It would be better, Tuchkevich urged, if the research institute could charge industry for what it has produced, return the initial investment to the statae, and use the profits for its own needs.

Vera Rich

Japanese physics

Accelerator swings into action

Tokyo

JAPAN's High Energy Physics Laboratory, KEK, has achieved a world's first by boosting the electron beam in the 3-km main storage ring of TRISTAN to 25.5 GeV during a test run of the accelerator. If tests of the positron beam go according to plan, the first collision experiments will begin before the end of the year and TRISTAN can enjoy a spell as the world's number one electron–positron collider. But plans for electron–proton collisions have been dropped in the face of competition from West Germany's HERA electron–proton collider.

TRISTAN, which has taken five years and Y87,000 million (£388 million) to build, consists of a 2.5-GeV linear accelerator that feeds electrons and positrons into an accumulator ring where they are boosted to 6.5–8 GeV before injection into the main ring. Radio-frequency cavities in four long, straight sections of the main ring accelerate the counter-rotating bunches of electrons and positrons which can be squeezed by superconducting magnets and smashed together in the middle of each straight. But as the electrons/positrons traverse the bends of the main ring near the speed of light, energy is released as synchrotron radiation, a loss which limits the power of the accelerator.

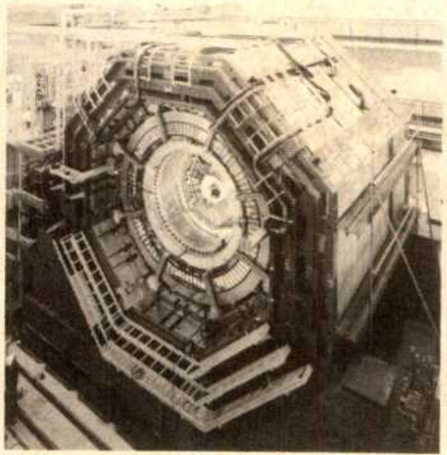
The record just set by TRISTAN exceeds the maximum of 23.5 GeV achieved by the PETRA ring in Hamburg, although TRISTAN's electron-beam current was only 0.4 mA. The KEK scientists are racing to raise this to 2–3 mA, the minimum required for collision experiments. VENUS, the first of TRISTAN's four detectors, is "about 70 per cent complete" and the first electron–positron collisions may be performed on schedule at the end of this month, according to KEK director Tetsuji Nishikawa.

Commissioning of TRISTAN ahead of US and European rivals will be a major triumph for Japan's high-energy physicists who have been struggling to put Japan back on the map since the end of the Second World War when the country's four largest cyclotrons were dumped at sea by US marines because someone in Washington feared they could be used to make atomic bombs.

But TRISTAN's dominance may be short-lived. The Stanford Linear Collider (SLC) has already taken electrons up to 37.5 GeV towards its target of 52 GeV with the prospect of electron–positron collisions in February or March 1987.

Nishikawa, however, is confident that with its higher luminosity and wider, stable beams TRISTAN will have advantages over SLC — the beams of SLC are only about a micrometre across and there will be difficulty in making them collide.

Among the particles TRISTAN may find are the top quark, Higgs particle, supersymmetric particles and fourth generation leptons and quarks, if they exist. But KEK scientists will also be looking for "totally unexpected and new phenomena in the new energy frontier".



TRISTAN will continue to consume large amounts of money, mainly in the form of electricity bills. In this fiscal year (until April 1987), Y6,000 million (£27 million) has been allotted by the Ministry of Education, Science and Culture for operation costs, enough for 1,000 hours. KEK is hoping for 1,800 hours at a cost of Y10,000 million in fiscal year 1987.

Next March, 40 radio-frequency cavities will be added to the main ring, bringing the beam energy up to about 28 GeV, and the addition of superconducting cavities in 1988 is expected to boost this to 33 GeV, a matter of months before the LEP electron–positron collider ring at CERN is due to come on line at 50 GeV.

KEK also plans to use the main ring as a 10-GeV synchrotron radiation source, which could be another first for Japan, but the most ambitious plan of all for TRISTAN has been abandoned.

Since TRISTAN was first proposed in 1973, it had been intended to include a proton storage ring alongside the electron–positron ring to allow electrons at 30 GeV to collide with protons of 300 GeV. Space for the proton ring was made in the 3-km tunnel of the main ring and a 12-GeV proton synchrotron built at KEK in 1976 would have provided the protons. But West Germany's decision to build the larger HERA electron–positron collider at Hamburg by 1990 has apparently caused Japan to change its plans. The main reason, says Nishikawa, is that HERA, which will smash 30-GeV electrons against protons of at least 820 GeV (with even higher energies in prospect), will be much more likely to reveal quark structure.

David Swinbanks

One-way crossing

THE British electrical utility, the Central Electricity Generating Board, began last Thursday to draw some 1,500 MW of cut-price electricity from France along four pairs of high-power cables recently laid across the English Channel. The French generating system is now more than 70 per cent nuclear and has over-capacity. Under the 1–2-year deal between CEGB and its opposite number, Electricité de France, the British board will take at least 1,000 MW continuously at a cost 25 per cent below 'average' generating costs in Britain, and 'part' (typically 500 MW according to CEGB) of the remaining 1,000 MW capacity of the link at the same cost. The link is two-way, so British electrical exports to France are possible, but none is planned. The power thus to be drawn by Britain from France exceeds the capacity of a single pressurized water reactor, such as that CEGB wishes to build at Sizewell in Suffolk, and on which the planning enquiry report is now imminent.

Robert Walgate

What anniversary?

THE BBC (British Broadcasting Corporation) is this week celebrating the fiftieth anniversary of its first television broadcast with a series of past performances almost calculated to meet the Conservative (and government) party's charge than the BBC's news reporting is biased; past politicians are mocked with an even hand. But readers should judge for themselves, from this extract from *Nature* (**78**, 151; 1908), whether the BBC is celebrating the right anniversary. □

Distant Electric Vision.

REFERRING to Mr. Shelford Bidwell's illuminating communication on this subject published in NATURE of June 4, may I point out that though, as stated by Mr. Bidwell, it is wildly impracticable to effect even 160,000 synchronised operations per second by ordinary mechanical means, this part of the problem of obtaining distant electric vision can probably be solved by the employment of two beams of kathode ray (one at the transmitting and one at the receiving station) synchronously deflected by the varying fields of two electromagnets placed at right angles to one another and energised by two alternating electric currents of widely different frequencies, so that the moving extremities of the two beams are caused to sweep synchronously over the whole of the required surfaces within the one-tenth of a second necessary to take advantage of visual persistence.

Indeed, so far as the receiving apparatus is concerned, the moving kathode beam has only to be arranged to impinge on a sufficiently sensitive fluorescent screen, and given suitable variations in its intensity, to obtain the desired result.

The real difficulties lie in devising an efficient transmitter which, under the influence of light and shade, shall sufficiently vary the transmitted electric current so as to produce the necessary alterations in the intensity of the kathode beam of the receiver, and further in making this transmitter sufficiently rapid in its action to respond to the 160,000 variations per second that are necessary as a minimum.

Possibly no photoelectric phenomenon at present known will provide what is required in this respect, but should something suitable be discovered, distant electric vision will, I think, come within the region of possibility.

A. A. CAMPBELL SWINTON.
66 Victoria Street, London, S.W., June 12.

European research

Applicants galore, but few grants

Four out of five applicants for research grants from the European Commission in Brussels fail to get a single ECU, although many of the project proposals are worthy of support, a Commission official warned British hopefuls in London last week. The occasion was organized by the Cabinet Office to give a clearer indication to British scientists and industrialists of the support available from Brussels and how to apply for it.

Apart from the shortfall between good applications and the support available, of which the Commission repeatedly complains when seeking more cash from the Council of Ministers, there can be long delays before decisions are reached. Successful researchers often receive the cash they need only in the year after they have spent it. The Commission admits its deficiencies and has set in train a programme of 'acceleration and simplification' to tidy up its act.

According to Commission administrator Roderick Hurst, the new 'framework programme' for Commission research and development now before the Council of Ministers will also improve matters by introducing majority voting on individual research programmes, so speeding up the time between a Commission proposal and final approval by the Council of Ministers.

Hurst says that the most frequent criticisms are of the lack of transparency in decision-making; the complexity of contract conditions (drawn up by Commission lawyers to satisfy the laws of 12 countries); the uncertain timing of calls for proposals (offering grants in specific areas); and the delays in processing and deciding to support particular proposals.

Announcements are made in the *Official Journal* of the European Communities ("which of course you all read"), but also through national Community offices, through government departments and by advertisements in journals. But one recipient of a grant says that the academic grapevine is the best means of communication.

There were a few tips for applicants last week. Nobody in paper-submerged Brussels has time to read 100-page theses, but the brief one-and-a-half page synopsis that applicants are asked to supply is critical, the sole basis for the first weeding out. Because referees may not be fluent in the language in which the synopsis is written, it must be exceptionally clear, but to the Commission it is now also essential that applications involve a collaboration among institutions from two different countries, preferably of different types (for example university and industry).

But having too many partners is impractical, listeners were advised; the usual range is three to five. Applicants, like the good examinees they no doubt often berate for the same fault, must also answer the questions on the form. "It is amazing how many people leave questions unanswered or answer them indirectly", Hurst said. Applications in non-nuclear energy are most likely to succeed (one in three accepted), followed by environment (one in four) and industrial technology (one in six). Biotechnology is at the bottom of the list, with a high demand and a low budget (only one in seven proposals in this field have been accepted).

The British Cabinet Office is preparing a booklet to assist potential applicants. According to a Cabinet Office spokesman, it is "no secret" that Britain was unhappy with the framework programme as presented to ministers earlier this year. The original documentation and legal instruments had not been a satisfactory basis for discussion, he said, though improvements are now hand. The central issue seems to be the British government's wish to apply to the Commission's proposals the disciplines it applies to its own expenditure, he said.

Robert Walgate

• The commission's presentation has been given at Leeds and Cardiff in the past week and is to be taken to Birmingham (12 November), Belfast (13 November) and Glasgow (15 November).

British research

Postdocs over the hill at thirty?

Prejudice in Britain against older researchers is widespread, dissuading many talented people from embarking on postdoctoral research for fear of being ruthlessly discarded within ten years. This is one conclusion from a study by the Association of Researchers in Medicine and Science (ARMS) based on a survey of applicants for recently advertised postdoctoral fellowships.

ARMS says that its survey reveals an unjustifiable prejudice in favour of people younger than 33 years. But far from attracting youth, the prejudice has caused an exodus of British talent abroad, or to other forms of research. The advertisers of the positions were generally disappointed with the number and the quality of the applicants.

The survey covered 200 postdoctoral positions, mostly in biology and medicine, advertised in the principal professional journals during three months from December 1985. ARMS says that fewer than 7 per cent of the positions were for longer than three years. The average duration of the posts on offer was 2.5 years and the average age of those appointed 28.3 years.

Most of all, the association is disturbed by the age distribution of the successful candidates. About a quarter were between 29 and 33 years of age, but only 6 per cent were older than 33. In eight cases, employers complained that many applicants were too old — in their thirties.

The association challenges this preference for younger researchers. It claims that postdoctoral scientists in their thirties are as intelligent, hardworking and competent as in their twenties, but with the advantage of greater knowledge and experience. That they should find themselves trapped in so uncertain and precarious a way of life ironically makes the whole research profession unattractive even to the younger people employers say they wish to attract, according to ARMS.

The report of the survey includes an account of the reasons given by advertisers of the posts for the few high-quality applicants. The common theme is that better long-term treatment has attracted more researchers to industry.

One, a biochemist from Leeds, said that "anybody in molecular biology who is any good has gone abroad, into industry or has left academic circles ... it is almost impossible to fill university research posts now". Another biochemist, based in London, was astounded at the poor response. Three years ago, he said, a similar advertisement would have attracted more than 40 good applicants, but none of the 13 applicants for the post advertised this year was suitable, and it had to be re-advertised.

A Dundee physiologist endorses this view. The quality of applicants for short-term posts has fallen because career prospects have dwindled. Financial rewards are poor compared with those in industry, and better prospects await those who emigrate to the United States, Europe and Australia.

Professor Peter Campbell, deputy chairman of ARMS and professor of biochemistry at Middlesex Hospital Medical School, has been trying to persuade those who support research to change their attitudes. Campbell argues that research often depends on teams of individuals with a range of abilities and spanning many disciplines. He thinks it proper that such teams should include both people with considerable originality and those with special techniques acquired over many years. But those of the second kind tend to be discarded by the system. Campbell also blames the salary structure in British universities. Because salaries are related to age, he says, researchers often find themselves priced out of posts in their mid-thirties.

Bill Johnstone

US cancer

Can NCI cut deaths in half?

Washington

CUTTING cancer deaths in half by the year 2000 is an achievable goal, according to the National Cancer Institute (NCI) In a monograph released last week*, NCI lays out its strategy for achieving that goal But its critics worry that the plans may do more to maintain the health of NCI than they will to reduce cancer mortality

Edward Sondik, coeditor of the monograph, says the NCI approach has two primary focuses, reducing smoking and improving treatment In 1985, 462,000 people in the United States died of cancer, and NCI reckons 30 per cent of those deaths are attributable to tobacco By 2000, NCI hopes to reduce cancer mortality attributable to smoking by between 8 and 16 per cent, implying that the number of people who smoke must be reduced to a half within this decade Although the percentage of the US population that smokes has been declining since 1965, the rate of decline would have to double for males and considerably more than double for females for this goal to be achieved

Reduction in mortality through treatment accounts for half the targeted decline Using new communication tools such as the Physician Data Query (PDQ), a computerized information service, NCI hopes modern therapies will be more widely disseminated NCI is also relying on regional cancer treatment centres and cooperating groups of physicians to test and implement new therapeutic regimens But Sondik concedes that more than half of the projected improvements attributable to treatment must come from treatments yet to be proved effective

In addition to smoking and therapy, the NCI report calls for changes in diet as well as more screening programmes for early detection Reducing fat to 30 per cent of total calorie intake and increasing fibre to 20–30 grams per day will yield an estimated 8 per cent reduction in mortality by 2000, according to the NCI report Another 3 per cent reduction in mortality may be expected from wider use of Pap smears for all women over 20 and mammograms for women over 50

But critics feel the NCI figures are overly optimistic John Cairns of the Harvard School of Public Health says NCI cannot justify the expected decreased mortality projected by the report With cancer deaths currently around 450,000 annually, Cairns says NCI projections require that more than 100,000 deaths be prevented annually through treatment Apart from breast cancer, where improved mortality rates have been achieved, saving approximately 10,000 lives annually, Cairns wonders where the other 90,000 or more will come from

A debate is currently raging over whether NCI is winning the war on cancer (see the *News and Views* article by Jared Diamond in *Nature* **323**, 488, 1986) which was officially declared in 1971, Sondik says improved long-term survival rates show that there have been significant strides in treatment for many types of cancer But others argue that age-adjusted mortality statistics paint a less encouraging picture and that a dramatic turnaround will be needed to achieve NCI goals

At the heart of the debate is whether NCI is receiving too large a share of the federal research budget, and whether that share is being spent properly Critics say it has put too great an emphasis on treatment at the expense of prevention, but Sondik counters that prevention efforts currently account for about 28 per cent of NCI outlays Treatment programmes have been receiving approximately 30 per cent of the NCI budget

What cannot be gainsaid is that the NCI budget is large, and growing In 1987 NCI will receive approximately $1,400 million, slightly less that one quarter of the total budget for the National Institutes of Health NCI is seeking $1,700 million for 1988 But its associate director Peter Fischinger is not sure the fight against cancer will cost much more in the future He hopes that the mix of prevention, screening and treatment will achieve the desired results

NCI's goal of halving cancer mortality by 2000 is not new, having been first announced in April 1984 at a congressional hearing NCI has steadfastly maintained that dramatic reductions in cancer deaths are possible But officials now sound a note of caution Fischinger says the goals for 2000 are not testaments, only estimates It will be up to Congress to decide whether those estimates are worth the hefty price tag **Joseph Palca**

**Cancer Control Objectives for the Nation 1985 – 2000* NCI Monograph No 2 1986

French research

Teetering on the brink

YOUNG French scientists who were offered jobs for life earlier this year by the leading French research council, the Centre National de la Recherche Scientifique (CNRS), but have since been denied this privilege, are taking their case to law

The 522 people concerned were appointed by the usual procedures in the annual recruitment to the research council, but then were effectively dismissed by a government decree invalidating the appointing committees The group accordingly formed a pressure group, the "Collectif des Admissibles", and appointed a lawyer to argue its case Legal proceedings have now begun, in the highest court of the land, the Conseil d'État, against science minister Alain Devaquet, who issued the decree, and (in a lower court) against the director-general of the research council, Serge Feneuille

Members of the group are also seeking substantial damages in private actions for loss of livelihood, expenses in relocation and other inconveniences Meanwhile, the research council has offered one-year contracts to 285 of the 522

The background is complex and political The Conseil d'État ruled on 12 May that certain elections to the appointments committees of the research council had been irregular In the event, the necessary corrections could have been made when the committees were re-elected for 1987 But the arrival of the new government and, it is said, pressure from right-wing physicians discontented with the survival of CNRS is thought to have persuaded the government to take the ruling seriously Devaquet's decision to dissolve the CNRS appointments committees and to annul all appointments made this year seems not to have been strictly necessary in the wake of the Conseil d'État's ruling That is the interpretation that many French scientists are putting on the affair

But what effect will it have on French science? A new Comité National (the body of appointments committees) is to be elected next year Feneuille, who comes from industry, is said to want to make the stopgap short-term contracts a permanent feature of CNRS life

Next year's Comité National will have to consider which of the 285 short-term appointees from the original 522 'admissibles' should stay, there is no guarantee that all will Those thinking of joining the research council are also in a quandary will the 285 have priority, affecting the numbers of new staff to be appointed?

For the longer term, French researchers are having to come to terms with lower recruitment rates than previously expected, as well as lower budgets They will also be facing a new Comité National, which, under a new constitution, will include representatives from the whole of the university teaching community, and not just from researchers as at present In the harder sciences, where about 60 per cent of university staff are in research and in some way connected with the research council, the effects of this change may not be too great But in medicine, where only 30 per cent do research, the new powers in control at CNRS may be evident

Robert Walgate

A cosmic rod in pickle

SIR—Under this title, Hoyle and Wickramasinghe[1] have tried to defend their various positions, and the claim that their earlier predictions had "wilfully and even maliciously been overlooked"

It seems clear, however, that the real reason that their work is not widely accepted or referred to is not that they have "sinned mortally" by making a "contact between biology and astronomy", but because so many of their various claims, put forward with so much certainty on the basis of unimpressive and even inaccurate evidence, have routinely been abandoned by themselves or proved wrong by others (see refs 2 and 3) Indeed, even before publication, their *Nature* article had been rendered obsolete by a published paper[4] they had submitted on 29 November 1985, but did not refer to This paper claimed hollow, spherical C_{60} (buckminsterfullerene) molecules[5] as the explanation of the composition and ultraviolet (UV) spectrum of interstellar grains that had been earlier identified, *inter alia*, as dried hollow bacteria[6,7] If the absorbence in the infrared (IR) were really due to such bacteria the absorbence in the UV would be far greater than in the IR[8,9] They cannot have it both ways, especially as the UV and IR spectra of this C_{60} molecule have not yet been measured (R F Curl, personal communication)

But there are other problems As it was part of their Fig 1, Karim, Hoyle and Wickramasinghe had to have known that the rest of the absorption spectrum of tryptophan did not fit the curves in their Figs 3 and 5 when they claimed that "the discovery of a broad interstellar absorption feature centred on the λ 2,800Å in the extinction curve of starlight confirms the presence of proteinaceous material in grains"[10] In ref 12, they must have deliberately cut off the low wavelength part of their own (mislabelled) curve in Fig 1 to make it fit the curves in Fig 4 to support one of their many now abandoned theories (For further comment, see refs 2–4, 8 and 11) Also, Hoyle *et al* [6] knew that my colleagues and I[2] had never claimed the existence of a peak for *Escherichia coli* at 220 nm, and so on This type of scientific behaviour is clearly unacceptable

In the past few years, many comments on their publications have appeared For example, Greenberg[13], McLachlan and Nandy[14,15] and Savage and Sitko[16] have shown that the claimed[10] absorption feature at λ = 2,800Å was produced by saturation effects in overexposed spectra and was flagged as such in the data bank

Thus, the long history of unwarranted claims by Hoyle and Wickramasinghe means that I, and many other scientists, will not accept anything they claim, be it about darwinism, viruses, bacteria, *Archaeopteryx*, or even interstellar grains or steady-state cosmology, unless and until it is demonstrated by others with far better batting averages and about whose works it cannot be said "That's not cricket" By the way, astronomers and chemists, also, have "rods in pickle"* for them too

R E DAVIES

Departments of Animal Biology, and
Astronomy and Astrophysics,
University of Pennsylvania,
Philadelphia, Pennsylvania, 19104, USA

1 Hoyle, F & Wickramasinghe, N C *Nature*, **332**, 509–511 (1986)
2 Davies, R E , Delluva, A M & Koch, R H *Nature* **311**, 748–750 (1984)
3 Davies, R E *Panspermia Unlikely, Unsupported, but Just Possible* Preprint 36th Congress of the International Astronautical Federation Stockholm, Sweden October 7–12 (1985) Int Astro Fed (Pergamon, Oxford)
4 Hoyle, F & Wickramasinghe N C *Astrophys Space Sci* , **122,** 181–184 (1986)
5 Kroto, H W , Heath, J R , O'Brien, S C , Curl, R F & Smalley, R E *Nature* **318,** 162–163 (1985)
6 Hoyle F , Wickramasinghe, N C & Al-Mufti, S *Astrophys Space Sci* **98** 343–352 (1984)
7 Hoyle, F , Wickramasinghe, N C & Al-Mufti, S *Astro phys Space Sci* **110,** 401–404 (1985)
8 Davies, R E , Delluva, A M & Koch, R H in *The Search for Extra terrestrial Life — Recent Developments Int astr Un symp 112* (ed Papagiannis M D) 165–169 (Reidel, Dordrecht, 1985)
9 Yabushita S , Wada K , Takai T , Inagaki, T , Young, D & Arakawa, E T *Astrophys Space Sci* **124,** 377–388 (1986)
10 Karim I M , Hoyle, F & Wickramasinghe, N C *Astro phys Space Sci* **94,** 223-229 (1983)
11 Koch, R H & Davies, R E *Astrophys Space Sci* **100,** 425–426 (1984)
12 Karim L M , Hoyle, F & Wickramasinghe, N C *Astrophys Space Sci* **100,** 431–435 (1984)
13 Greenberg J M *Observatory* **104,** 134–135 (1984)
14 McLachlan, A & Nandy, K *Observatory* **104,** 29–31 (1984)
15 McLachlan, A & Nandy K *Astrophys Space Sci* **109,** 399–402 (1985)
16 Savage, B D & Sitko, M L *Astrophys Space Sci* **100,** 427–429 (1984)
19 Partridge, E *A Dictionary of Slang and Unconventional English* 7th edn (Macmillan, New York, 1972)

*For those not familiar with this somewhat obscure term, a "rod in pickle' , which Hoyle claims to have "not a few of", does not mean a venereally infected penis (UK) a good horse that is being either reserved or nursed for a sure win (Australia), a pistol in a plight (USA) but a prospective punishment or scolding (UK)[18]

Indian collaboration

SIR—The news item "Indian science Foreign software giants pounce" (*Nature* **323,** 102, 1986) states "Astra's subsidiary in Bangalore has taken the cream of Indian biochemists, including the chief biotechnologist of the Indian Institute of Science" As I am a biochemist who has acquired the label "biotechnologist" and am intimately involved in the organization of the Astra centre, I wish to say that the statement is not correct I am not on the payroll of Astra nor have any of my colleagues left the institute to join Astra The core group at the centre consists of fresh PhDs who were ready to leave the country and take up postdoctoral fellowships in the United States Besides, Astra is trying to bring back qualified Indians from abroad, so the centre is helping to prevent a brain-drain, both internal and external

Astra has given funds to the Indian Institute of Science, enabling me to further my basic research, which I have been carrying out for several years At the same time, we have an opportunity to interact with the centre on problems related to the alleviation of human suffering The Astra centre seems to us to be a unique, open and fair collaboration between an academic and an industrial research institution The Astra centre in Bangalore is not a subsidiary of Astra Sweden The centre will develop its own know-how and is recognized by the government as a non-profit research organization It is registered as a society in Karnataka State The government of India has laid down conditions to protect the Indian interests in terms of publications, patents, royalties and the conduct of research

There is much talk about the virtues of collaboration between industry and academic research There is also criticism that scientists in India have not been doing relevant research and that they should have discovered cures for tropical diseases a long time ago Now that scientists are trying out a new model of collaboration and when the government has done what it can to protect Indian interests, it is a pity to lump together and to condemn all multinational efforts I wish K S Jayaraman had made a deeper study of this collaboration model before offering his comments or those of his critics, which could harm not only the institution but the entire philosophy

G PADMANABAN

Department of Biochemistry,
Indian Institute of Science,
Bangalore 560 012, India

Faraway quasars

SIR—You quote me (*Nature* **323,** 193, 1986) as saying that it is only a matter of time before very long baseline interferometry yields the absolute distances to at least one quasar with an interposing gravitational lens Alas, were that only true What I said, in fact, was that we know of no way to determine directly the distances to quasars But if a gravitational lens between the quasar and us forms multiple images, then, if the quasar cooperates, for example by changing its luminosity sharply, it is possible, at least in principle, to determine the *differences* in the times taken by light to propagate from the quasar to us along the paths for the different images I stated my opinion that it would just be a matter of time before an imaged quasar cooperated sufficiently for us to determine such time differences

IRWIN SHAPIRO

Harvard-Smithsonian Center
for Astrophysics,
60 Garden Street,
Cambridge, Massachusetts 02138, USA

Physicists about to hi-jack DNA?

Unless molecular biologists are careful and energetic, they will find that the mechanical and thermodynamic properties of DNA have been taken over by physicists.

MOLECULAR biologists, if pressed, will readily agree that more should be done to throw light on the properties of naturally occurring macromolecules in the molecular environment in which they ordinarily exercise their biological activity, usually that of water. More than that, much is being done, experimentally and otherwise. X-ray diffraction measurements of, say, crystalline proteins are not always the abstract studies they may seem, given the large water content of many specimens. The application of nuclear magnetic resonance techniques to enzyme molecules may indirectly say much about water interactions. The abiding obstacle, in these connections and others, is that the important molecules are so often complicated in their own right, even without the extra complication of their external interactions. People sigh for simpler systems.

That is where physics may have the edge. There is probably a great deal in the opinion that biologists are reconciled to having to tackle the problems with which nature confronts them, but that physicists are free to pick and choose, tackling only the problems they can reasonably hope to solve. Put starkly, like that, the distinction is obviously a simplification (Sir Peter Medawar's *The Art of the Soluble* is, in any case, a biologist's book.) But the caricature does seem to illustrate a difference of style, which is in turn nicely illustrated by a neat analysis of the properties of DNA molecules in solution by L. L. Van Zandt of Purdue University appearing in *Physical Review Letters* (**57**, 2085, 1986).

The starting point for Van Zandt's analysis is an account which appeared two years ago (Edwards, G. S., Davis, C. C., Saffer, J. D. & Swicord, M. L. *Phys. Rev. Lett.* **53**, 1284, 1984) of the absorption of microwaves in solutions of DNA. Part of the motivation of that work was the belief that measurements might throw light on suggestions that microwave absorption can cause biological damage even when the amounts of energy involved are very small, perhaps because of some resonance between the microwave frequency and a vibrational frequency of some vital and susceptible molecule.

But how could such a molecule, say a DNA molecule, even if characterized by well-defined frequencies of vibration, accumulate substantial amounts of energy in these vibrational modes while surrounded by water molecules whose effect would be that vibrations of any kind would be efficiently damped? Edwards and his colleagues were sustained by their earlier observation of a dramatic increase in the gross absorption of microwaves by solutions of DNA from *Escherichia coli* in the presence of the endonuclease DNase I. They inferred that the continuing degradation of long molecules would assure a constant supply of shorter molecules capable of vibrating resonantly with external microwaves.

So why not test this possibility by using the techniques of genetic manipulation to make pieces of DNA whose size is well defined? Edwards and colleagues did just this and made two remarkable discoveries. First, well-defined DNA molecules, whether circular or linear, do indeed show resonance absorption of microwaves in the frequency range from 1 to 10 GHz. Second, the absorption peaks are surprisingly sharp, suggesting ineffective damping by the surrounding medium.

Why that should be is for Van Zandt to say, but the numbers are interesting in themselves. The measured damping implies that the relaxation time, that needed for a vibrational impulse to be dissipated, is of the order of 300 picoseconds. But the velocity of sound in DNA is within 15 per cent of 2 km s^{-1}. One consequence is that vibrational excitations can travel only about 1,500 base pairs along a linear DNA molecule before being dissipated, which in turn implies that coherent vibrations over much greater distances will not easily happen. That is why Edwards *et al.* found that their resonance frequencies were usually harmonics of a fundamental.

Van Zandt has had the benefit of an even more striking demonstration of anomalous resonance in DNA (by the Edwards group) based on measurements of DNA molecules with 5,480 base pairs derived from a circular plasmid in which supercoiling has been relieved by the use of a topoisomerase. The remarkable feature of the measurements is that the strength of the absorption (four peaks between 3 and 9 GHz) increases with increasing frequency, the opposite of that expected of an efficiently damped system.

The explanation can be made simple enough to suit any physicist. First, the vibrations responsible for the resonances are simple organ-pipe vibrations — standing compressional waves set up around the circular channels of the plasmids. A few simple properties of the solvent, water, are relevant. Energy dissipation in pure water arises because of the stretching and distortion of hydrogen bonds, but because these form and reform in times of the order of 10^{-11} seconds, they should be measurable only at frequencies of 50 GHz and above. Van Zandt's question is how circular DNA can so profoundly modify this behaviour.

The molecular model is simple. Suppose that there is a single layer of water molecules more tightly bound to the DNA than to the bulk of the solvent, so that it will move with the molecule if that should vibrate, but at the same time be better adapted than DNA to sense and be influenced by the molecular motions in bulk liquid. Van Zandt constructs the crudest of all possible mechanical models to represent this state of affairs. Take two masses, one representing the mass of DNA (per unit length) and the other the mass of the bound water layer (again per unit length) and imagine them connected by an elastic spring. Let the second mass (the bound layer of water molecules) be damped, and calculate the absorption of the system as a function of microwave frequency.

Sceptics will say that even this simple system has enough adjustable constants to fit pretty well any measured curve. Van Zandt plays fair by using real values (the mass of DNA and the bulk viscosity of water) to fit as many of them as possible. Even the allowable frequencies of vibration are taken from measured velocities of sound in dry DNA (and the condition that there must be an integral number of wavelengths around a single circle of DNA). The result is a remarkably good fit of the position of the four absorption peaks, but the measured peaks are much broader than predicted.

Where this will lead is anybody's guess. Van Zandt promises a more detailed account of the structure of the solvent. The discrepancy between the width of the measured and predicted peaks cannot be made to go away by adjusting (and increasing) the damping constants in the model, for that has the effect of changing the relative strength of the peaks. Van Zandt thinks that even a circular plasmid whose supercoiling has been relieved may have some residual tertiary structure, but there are others who will see the difference as an opportunity to learn more about the solvent interactions. How soon will it be before molecular biology laboratories equip themselves with a klystron and some waveguide?

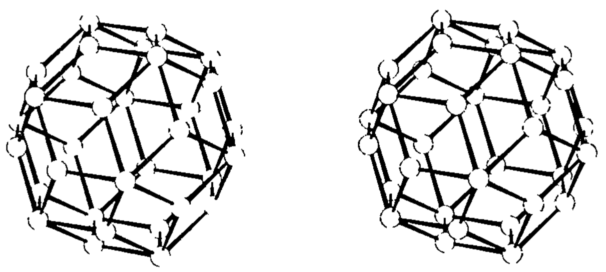
John Maddox

Five-fold symmetry

Millimetre-sized quasicrystals grown from aluminium alloy

Kevin Knowles

QUASICRYSTALS, materials with 'forbidden' five-fold symmetry, were discovered two years ago, but their investigation has been dogged by the lack of large crystals. The critical X-ray and electron diffraction data which demonstrate five-fold symmetry are from very small quasicrystal particles, complicating interpretation of diffraction intensities which so far do not seem to be consistent with any clear atomic arrangement. But investigations should take a leap forward with the report of Dubost *et al* on page 48 of this issue[1], which describes the growing of quasicrystals approaching a millimetre across, and raises the real prospect that clear X-ray and neutron diffraction pictures of single quasicrystals will now be possible.



Rhombic triacontahedron, arranged as a stero pair for a three-dimensional view (courtesy of A. Mackay, University of London)

One of the basic principles of morphological crystallography is that the external shapes of crystals are determined by the point-group symmetry of the underlying fundamental units of the material, the unit cells, which are typically of nanometre dimensions. Thus most scientists in their schooldays will have grown crystals of alum from solution and will have found that although some will crystallize in regular octahedra characteristic of the underlying cubic symmetry, most look irregular, but nevertheless can be described in terms of simple arrangements of the unit cell.

The quasicrystals of Dubost *et al* are grains of an Al–Li–Cu alloy showing triacontahedral morphology. They are 'quasicrystalline' in the sense that their morphology alone would imply an underlying 532 or m$\bar{3}\bar{5}$ point-group symmetry, incompatible with translational symmetry inherent in the perfect crystalline state. The grains grew from the 'T2' phase of Al–Li–Cu, one of the family of icosahedral phases whose structure and properties have been of intense experimental and theoretical interest since the initial reports of these phases in alloys of Al–Mn, Al–Fe and Al–Cr[2-4].

The T2 phase is not only of academic interest — it occurs as a stable grain boundary precipitate in commercial Al–Li–Cu–X alloys such as those now being developed by aluminium manufacturers for light aerospace materials with consequent savings in fuel bills. The dimensions of the Al–Li–Cu T2 grains grown by Dubost *et al* under conditions of very slow cooling, allowing only a few nuclei to grow, contrast with the precipitates of micrometre or sub-micrometre dimensions produced from other icosahedral phases, for which the cooling rates needed to produce the phases are typically of the order of 10^7 K s^{-1}, similar to the fastest quench rates reported for obtaining crystalline phases, but far slower than quench rates for metallic glass formation. Indeed, these T2 grains are sufficiently large to produce X-ray and neutron diffraction patterns from single crystals. The results from such experiments are keenly awaited, as the size of quasicrystals produced to date has allowed only powder X-ray diffraction data to be obtained.

These results should help to determine the origin of the line broadening in powder X-ray diffraction patterns that can be interpreted as evidence of either frozen-in phason strain or random icosahedral packing[5,6]. (The simplest way to see a phason in quasicrystalline materials is as a local violation in the matching rules in a Penrose tiling[7], switching, for example, the internal vertices of rhombic triacontahedral units within the tiling.)

The ability to produce large grains of icosahedral phase will also be important for other electron-beam work, which indicates that the chemical composition of icosahedral phases grown under rapid cooling conditions is inhomogeneous on a 5–10-nm scale. Chen *et al*[8] have reported changes in the concentration of Mn in Al–Mn icosahedral phase from place to place, and dark-field transmission electron microscopy reveals cloudy and speckly contrast that can vary as a function of diffraction vector, beam direction and alloy formation[9-14].

This work implies that structural features analogous to spinodal decomposion or anti-phase domain boundaries in crystalline materials are inherent in rapidly solidified icosahedral phases. Will this also be found to be the case in the large T2 Al–Li–Cu grains where presumably equilibrium is more easily reached?

The mechanical properties of this phase are particularly relevant. Theoretical considerations by Lubensky *et al*[15] on the basis of hydrodynamic arguments predict that quasicrystals are inherently extremely brittle as dislocation motion is virtually inhibited by the slowness of phason modes in these materials.

On this basis, Lubensky *et al* suggested that it would be virtually impossible to fabricate and manufacture macroscopic quasicrystals (such as the ones grown by Dubost *et al*), although they recognized that the dynamics of dislocation motion in quasicrystals would alter radically if the phason modes were rendered non-hydrodynamic. It is therefore of interest to discover whether the T2 phase is indeed extremely brittle or whether it is instead merely similar to other intermetallic phases, which in general are characterized by high hardness and limited ductility. Dislocations, if they exist in icosahedral materials, must, on the basis of current electron microscope evidence, be either extremely rare or otherwise not exhibit the characteristic strainfield contrast seen in crystalline materials.

It seems clear that the report of Dubost *et al* will stimulate more work on icosahedral materials, particularly on the thermal stability and energetics of the phase formation in Al–Li–Cu. We can only imagine where our understanding of icosahedral materials would be if more than 30 years ago Hardy and Silcock[16], the original discovers of the T2 phase, had had an electron microscope and had been able to obtain similar icosahedral electron-diffraction patterns to those found by Shechtman *et al*[2] in Al–Mn. ☐

1. Dubost, B., Lang, J.-M., Tanaka, M., Samfort, P. & Audier, M. *Nature* **324**, 48 (1986).
2. Shechtman, D., Blech, I., Gratias, D. & Cahn, J. W. *Phys Rev Lett* **53**, 1951 (1984).
3. Mackay, A. L. & Kramer, P. *Nature* **316**, 17 (1985).
4. Mackay, A. L. *Nature* **323**, 293 (1986).
5. Horn, P. M., Malzfeldt, W., DiVincenzo, D. P., Toner, J. & Gambino R. *Phys Rev Lett* **57**, 1444 (1986).
6. Stephens, P. W. & Goldman, A. I. *Phys Rev Lett* **56**, 1168 (1986).
7. Socolar, J. E. S. *J. de Physique* **47**, Coll. C3 p 217 (1986).
8. Chen, C. H., Joy, D. C., Chen, H. S. & Hauser, J. J. *Phys Rev Lett* **57**, 743 (1986).
9. Shechtman, D. & Blech, I. A. *Met Trans* **A16**, 1005 (1985).
10. Field, R. D. & Fraser, H. L. *Mat Sci Engng* **68**, L17 (1984–85).
11. Urban, K. *et al J. de Physique* **47**, Coll. C3 P 465 (1986).
12. Knowles, K. M. & Stobbs, W. M. *Nature* **323**, 313 (1986).
13. Saito, Y., Takemoto, A., Tanaka, T. & Mihama, K. *Proc XIth Int Cong Electron Micros* p 523 (1986).
14. Robertson, J. L., Misenheimer, M. E., Moss, S. C. & Bendersky, L. A. *Acta Metallurgica* **34**, 2177 (1986).
15. Lubensky, T. C., Ramaswamy, S. & Toner, J. *Phys Rev* B32, 7444, 33, 7715 (1986).
16. Hardy, H. K. & Silcock, J. M., *J. Inst Metals* **84**, 423 (1955–56).

Kevin Knowles is in the Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB2 3QZ, UK

Molecular biology

Manipulation just off target

R Michael Liskay

PROGRESS has been made in several laboratories towards gene targeting — recombinational interactions between an introduced DNA molecule and its homologous sequence residing in a chromosome — in mammalian cell systems[1-4] The targeting studies of Thomas and Capecchi reported on page 34 of this issue[5] are provocative in several respects They indicate that DNA introduced into the nucleus by microinjection can induce mutations at high frequency in the target gene in a totally unexpected manner The observations both sound a note of caution for some applications of gene targeting and provide new methodologies that could facilitate the efficient introduction of mutations into a particular region of a gene

High-efficiency targeting is a potentially powerful approach for manipulating the genomes of higher organisms in a manner analogous to the molecular gymnastics that are possible in yeast, including the creation or removal of mutations in any chromosomal sequence for which one possesses a molecular clone

Thomas and Capecchi derived recipient mouse cell lines containing a mutated *neo* gene whose normal activity confers resistance to the toxic compound G418 This particular mutant form of the gene, here called the resident gene, is an 'amber' mutation that causes premature termination of translation and a resultant truncated protein The amber mutation (a single base-pair change) is located in the initial portion of the coding region of *neo* and its presence can easily be detected by molecular techniques Thomas and Capecchi attempted to correct the resident *neo* gene by microinjecting a different mutant form of this gene, a deletion that removes the last 52 amino acids from the protein, directly into the nucleus of the recipient cell At a frequency of approximately 1 per 1,000 injections they recovered cells resistant to G418 Roughly half of these lines now contained, as expected, an apparently normal *neo* gene — that is, the amber mutation had been removed Surprisingly, in the rest of the G418^r cells the amber mutation was maintained Eight of these exceptional, corrected genes, each representing an independent occurrence, were cloned into bacteria and their nucleotide sequences near the amber mutation were determined Each gene now harboured a compensating mutation Three different insertion mutations were found insertion of G, of T or of GGCT These insertions now allow a normally silent and out-of-phase protein synthesis start signal to be used A functional protein is produced that is incorrect for the first few amino acids but normal for the rest Furthermore, the mutations occurred within the first of two directly repeated sequences (GGCTAT) That each insertion results in a tandem duplication of at least a portion of the repeat sequence leads the authors to conclude that direct repeats are a prerequisite for this phenomenon

Control experiments, including microinjection of either the same amber mutant of *neo* or an unrelated DNA molecule, did not yield any G418^r cells From these control injections the authors conclude that the process requires both homology and at least one nucleotide difference, or mismatch, between the interacting genes Because of this apparent requirement for a mismatch, Thomas and Capecchi invoke the formation of heteroduplex DNA between the resident gene and the incoming plasmid molecule as an intermediate and therefore refer to this newly observed process as heteroduplex-induced mutagenesis They point out that because restoration of a functional gene was intrinsic to their experiments, many (even most) of the mutations induced by these interactions would be undetectable, making the actual frequency of mutations much higher than the 1 in 1,000 seen

Thomas and Capecchi also analyse how the compensating mutations result in functional protein These experiments provide a glimpse at how the ribosome can apparently reinitiate protein-chain elongation after encountering a single stop signal, provided there exists another start signal in the immediate vicinity

The main issue, however, is the impact that these findings have on various applications of and strategies for gene targeting in mammalian cells One might first ask why others studying targeting in mammalian cells have not observed error-prone events The answers seem to be relatively simple In one recent study[2] direct selection for targeted integration was not applied and there was no reason to examine the targeted gene at the nucleotide level In other experiments involving direct selection in a manner similar to that of Thomas and Capecchi, the nature and/or positions of the mutation in the resident sequence were such that the detection of error-prone events would often be precluded[1,3,4] For example, extensive terminal or internal deletions of the resident gene would not be prone to correction by the proposed mutagenic mechanism Note that in the bacterium *Salmonella* there exists a possibly related phenomenon termed selfing whereby one mutation can be rescued by the very same mutant DNA[6] Unfortunately, these events have not to my knowledge been analysed at the molecular level On the other hand, gene conversion, a form of nonreciprocal recombination, has been shown to act with fidelity in the several cases examined in fungi[7-9]

There are two general implications of heteroduplex-induced mutagenesis First, the results provide a caution to those who wish to manipulate mammalian genomes by gene targeting It is possible that unpredicatable and unwanted mutations could be introduced into a gene intended to be the target for correction For gene therapy in humans these potential side effects should be considered serious Second, and more positive, Thomas and Capecchi's observations suggest new methods for directly mutating mammalian chromosomal sequences that differ from the approaches that are standard fare in the yeast *Saccharomyces cerevisiae* In this yeast, specific mutations are engineered into a cloned sequence one-by-one and used to replace one-by-one the normal sequence by a targeted interaction Based on heteroduplex-induced mutagenesis, a single appropriately constructed plasmid might prove sufficient to induce a library of mutations in a particular region of the gene This approach might be used to advantage for mutating the control region (for example, promoter) of genes to define the location and function of *cis*-acting regulatory sequences Such regions frequently contain repeated sequence elements and therefore should be prime targets for the proposed mutagenesis by incoming DNA

One obstacle preventing the widespread use of targeting procedures is the propensity of mammalian cells to incorporate DNA into the genome in an essentially random fashion The efficiency of homologous relative to nonhomologous interactions is presently being achieved at a ratio of about 1 to 100 This ratio must be improved before the mammalian genome becomes as amenable to genetic tinkering as its smaller cousin, the yeast □

1 Smith, A J H & Berg, P *Cold Spring Harb Symp quant Biol* **49**, 171-181 (1984)
2 Smithies, O *et al Nature* **317**, 230-234 (1985)
3 Lin, F L , Spearle, K & Sternberg, N *Proc natn Acad Sci U S A* **82**, 1391-1395 (1985)
4 Thomas K R , Folger, K R & Capecchi, M R *Cell* **44**, 419-428 (1986)
5 Thomas, K R & Capecchi, M R *Nature* **324**, 34-38 (1986)
6 Demerec, M *Proc natn Acad Sci U S A* **48**, 1696-1704 (1962)
7 Roman, H *Cold Spring Harbor Symp quant Biol* **21**, 175-185 (1956)
8 Case, M & Giles, N *Genetics* **49**, 529-540 (1964)
9 Fogel, S & Mortimer, R *Molec gen Genet* **109**, 177-185 (1970)

R Michael Liskay is in the Departments of Therapeutic Radiology and Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510-8040, USA

Elementary particles

Colour theory in a spin

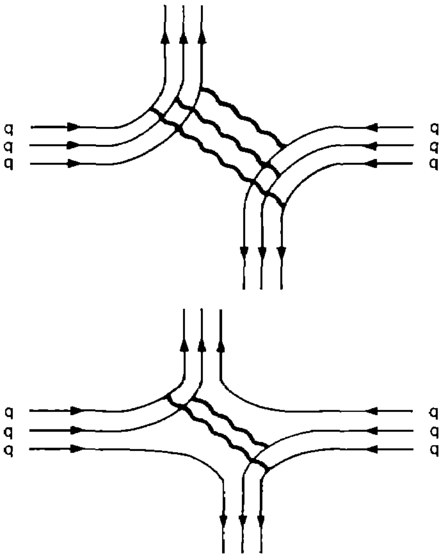
Harry J Lipkin

EXPERIMENTS in which beams of high-energy polarized particles are scattered by polarized protons[1-5] continue to puzzle, confuse and cause controversies among high-energy physicists. New experiments show a strong dependence of the scattering processes on the directions of the proton spins, which is not simply explained by accepted theoretical models based on quantum chromodynamics (QCD). The original establishment view was that little could be learned from polarization experiments because theory predicted no spin dependence at high energies. Subsequent investigations with QCD predicted some spin dependence, but not enough to explain the experiments. Thus some experimenters claim that they violate the predictions of QCD. The most recent experiment[1] suggests that this strong dependence suddenly disappears mysteriously at a higher energy. This causes even more confusion.

Spin physics was originally considered trivial at high energies, because of helicity conservation[6,7]. Particles with spins, such as electrons, neutrons, protons and quarks, are called right-handed or left-handed, respectively, if their spins are oriented parallel or antiparallel to their direction of motion. QCD predicts that this handedness (called helicity) is conserved at high energy, that is, it does not change in collisions at high energy, although it may do so at low. Thus, QCD seems to predict that protons emerge from any collision with their original helicity, and that no further useful information can be obtained from any spin measurements.

The argument that helicity conservation makes spin physics trivial is strictly valid only for elementary spinning particles. In the 1950s physicists believed that protons and neutrons were elementary particles and that Yukawa's meson theory of nuclear forces explained nuclear structure and the forces holding nuclei together, just as quantum electrodynamics (QED) explained atomic structure and the forces binding atoms together. Soon many new particles, now called hadrons, were discovered and it became clear that none was elementary. We now believe hadrons are built from smaller building blocks called quarks, bound together by QCD forces. The proton is no longer elementary but contains several quarks spinning in opposite directions. Quarks with opposite helicities can be exchanged between the two protons, thus changing the proton helicities even if the helicity of each quark is unchanged. Spin measurements therefore enable us to explore the internal helicity structure of the proton.

Spin effects played a crucial role at a very early stage[6-8] in providing experimental tests of the quark structure of hadrons. Already in 1945 the anomalous magnetic moment of the proton was interpreted as evidence that the proton was not an elementary particle like the electron, satisfying the Dirac equation. The quark model gave the first satisfactory explanation of the magnetic moments of the proton and neutron, by having these particles made from two kinds or 'flavours' of quarks, now called up and down, with electric charges +2/3 and −1/3 of the proton charge, respectively. The proton is made of two up quarks and one down quark, the neutron of two downs and one up, thus giving the observed charges of +1 and 0. Because all quarks have the same spin as the proton, the right total spin for the proton and neutron is obtained by having the up and down quarks spin in opposite directions, giving a total spin which is the difference between the up and down quark spins. But quarks with opposite electric charges spinning in opposite directions have their magnetic moments all pointing in the same direction and add up to a large value. This explains why the proton with a positive charge has a magnetic moment roughly three times the Dirac value and the neutron with no charge has a magnetic moment corresponding to roughly double the Dirac value for a particle with negative charge. A quantitative calculation of the ratio of the proton and neutron moments with this picture gives the value −3/2, in remarkable agreement with experimental results.



Two protons, each consisting of three quarks (q) collide at high energy and are scattered through a right angle. Top, direct scattering three high-energy gluon exchanges (wavy lines) between quarks in the two protons cause the deflection. Bottom, exchange scattering in each proton, internal motions have already produced a quark moving in the final direction, and only two high-energy gluon exchanges are required. The amplitudes of these diagrams are added to calculate the probability of the whole process.

The first excited states of the neutron and proton, now called Δ, had a spin three times the spin of the proton. These are now described as being made of the same three quarks with all spins parallel. There were also new companions of these particles, originally called strange particles. The greek letters Λ, Σ and Ξ denote the different strange particles with the same spin as the proton, an asterisk is added to denote the excited states. These are now believed to contain one or more quarks of a third flavour called strange quarks. In 1966 Federman, Rubinstein and Talmi[9], and independently Sakharov and Zeldovich[10,11], calculated the excitation energies of the parallel-spin states using an extremely simple model of the spin-dependent forces based on the analogy with hyperfine interactions in atomic physics. These authors obtained relations between the masses which were in excel-lent agreement with experiment, for example

$$\Delta - \text{proton} = (1/2)\,(2\Sigma^{*}+\Sigma-3\Lambda)$$

The experimental values of the two sides of this equation are 307 and 294 MeV, thus supporting the assumption that the spin forces between quarks are like the hyperfine interactions in atoms.

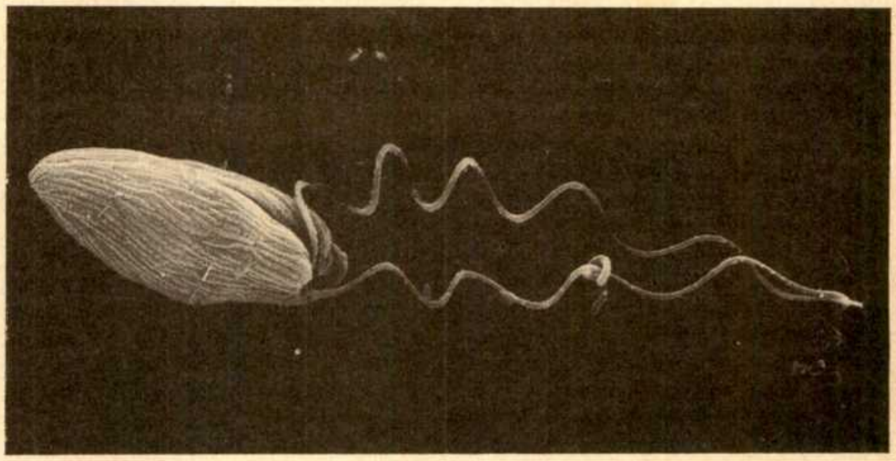
There is now an entire family of hadrons called baryons, all, like the proton and neutron, made from three quarks. There is also another family of hadrons called mesons, made of a quark and its antiparticle the antiquark. These include the pion (π) which has the two spins antiparallel, its strange partner the kaon (K) and their lowest excited states, denoted by ρ, ω and K*, with their spins parallel. Sakharov and Zeldovich[10,11] used their assumed hyperfine interactions to obtain a remarkable relation between meson and baryon masses

$$\Lambda - \text{proton} = (3/4)(K^{*}-\rho) + (1/4)(K-\pi)$$

The experimental values of the two sides are 177 and 180 MeV, giving striking support to the assumption that mesons and baryons are made of the same quarks.

There are many more examples of how spin effects have helped to reveal crucial clues to hadron structure which eventually led to QCD. I will now consider the attempts to understand spin effects in proton–proton scattering at high energies.

Perhaps some future article will begin "Back in the 1980s physicists believed that quarks were elementary particles and that QCD explained hadron structure and the



THIS unusual microorganism was recently found by Mr Harry Burton and Dr Ian Bayly in the course of an Australian Antarctic Division examination of a meromictic lake. These lakes have two well-defined liquid layers below the ice overlay; the top layer has a high oxygen content and low salinity, the lower layer has no oxygen but does have high concentrations of ammonia and hydrogen sulphide. The boundary between the layers is called the chemocline. The microorganism — currently known as Eric — is found only in the 30 cm of water directly below the chemoline. To go with this habitat, it has an unusual lifestyle; its outer surface is covered with bacteria, which appear to cling together with 'wing-like' structures. The bacteria have not been found separately. The little that is known about Eric is reported in *Science News, Fiji;* the Australian team are at present trying to breed from him.

Rebecca Ward

---

forces binding quarks together". These speculations that even quarks and leptons may be composite systems can be avoided by assuming that all hadron structure physics is described by QCD just as all condensed matter physics is described by QED. But no one knows how to calculate the properties of a type II superconductor from first principles with QED; similarly there is no clear recipe for the application of QCD to the description of hadron structure and the prediction of experimental results. In both cases the systems and the dynamical equations are too complicated. Theorists use QCD hand-waving arguments to pick a principal or 'leading' contribution for calculating experimental predictions. If a prediction agrees with experiment, they are happy and write a paper. If it disagrees with experiment, they write a paper anyway, and blame the disagreement on nonleading effects. If there are no experimental results available, they publish their result as a QCD prediction. If subsequent experiments disagree with this prediction, experimenters often claim that their results disagree with QCD, rather than with a particular choice of a leading term.

Experiments in the 1970s which scattered high-energy electrons from protons showed that at high momentum transfers the proton behaves like an assembly of free quarks with very weak interactions, even though the interactions between quarks are very strong at low energies. The fact that these interactions become much weaker at high energies has been given the name asymptotic freedom and shown to be predicted by QCD. QCD arguments suggest that experiments at sufficiently high energies can be treated with the use of the leading terms in an expansion called perturbative QCD, because the interaction has become weak enough to be treated by the same perturbation techniques used in QED. But no one has a convincing prescription for where high enough energies begin, and there is a running controversy on this point between different groups of theorists.

The recent polarization experiments are not explained by treatments using any of the obvious leading terms. The first experiments showing nontrivial spin effects motivated new calculations of polarization effects based on perturbative QCD suggesting that these experiments can give useful information after all about hadron structure and QCD dynamics. The simplest leading term predicted that the probability for proton–proton scattering by 90° should be twice as large if the spins of the protons in the beam and in the target were parallel than if they were antiparallel[12,13].

Theorists were delighted when new experiments approached just this factor of two[3] and seemed to level off there, but further experiments at higher energies showed a continued increase to factors as large as four[4]. New models arose, all paying lip-service to QCD, but with different leading terms. The latest result[4] further confuses the theorists by giving the same value for the scattering of protons with parallel and antiparallel spins, implying a sudden disappearance of the large difference which had risen to a factor of four at somewhat lower energies. There are also other experiments showing further disagreements with theory. Single-helicity-flip transitions were observed in 28-GeV proton–proton scattering at high transverse momentum[5], whereas the quark-exchange mechanism with helicity conservation at the quark level cannot produce a single helicity flip. An additional unexpected polarization effect was seen when ρ mesons were produced in collisions between pions and protons[14]. Strange particles producted in proton–proton collisions have long been known to be strongly polarized for reasons still not understood.

All these experiments suggest that there are interesting physics in these spin effects that are not understood in the simple leading terms and which may give clues to the underlying hadron structure and QCD dynamics. It is still an open question whether further experiments and analysis can reveal the relevant nonleading effects and lead to a better understanding. At present the only viable approach seems to be to continue the search for new clues in the data and try to find signals in the noise.

The basic physics and the essential complications of applying QCD to high-energy scattering can be seen by considering a colliding beam experiment in which two proton beams collide, one coming from the left and one coming from the right, and two proton detectors look for protons scattered at an angle of 90° with respect to the incident beams, one scattered upward, say, and one scattered downward. Because each proton consists of three quarks, the collision process begins with three quarks coming in from the left and three from the right, and it ends with three quarks going upward and three going downward. A theory based on QCD must explain how these six quarks get the appropriate kicks to turn their directions of motion by 90°.

There are several possible scenarios (see figure). The direct scattering scenario has all three quarks that came in from the left going out upward, and all three quarks that came in from the right going downward, or vice versa. The exchange scenario has two of the quarks that came in from the left going out upward and one going downward, and two of the quarks that came in from the right going downward and one going upward or vice versa. In the direct scenario, each of the three quarks coming from the left must have its momentum turned by 90° by exchanging one or more gluons with quarks coming in from the right. The exchange scenario makes use of the complicated internal motion of the quarks inside each proton. At the instant before the collision, there is a certain probability that this internal motion inside the proton coming in from the left will already have two of the three quarks moving upward and one moving downward, and that the proton coming in from the right will already have two of the three quarks moving downward and one moving upward. In this case it will be natural for the three quarks already moving upward before the collision to combine to make a proton going upward, and similarly for those moving downward.

Perturbative QCD suggests that as the interaction is weak at high momentum transfers, the scenario with the minimum number of high momentum kicks should win. The direct scenario requires high momentum kicks to turn each of the six quarks by 90°; this is very difficult at high energies. In the exchange scenario, the exchanged quarks are already going in nearly the right direction before the collision, the kick required is smaller and the QCD interaction stronger, so it is much easier to turn them around. However, the probability that they are initially moving in nearly the right direction is small. Thus two small factors enter into any theoretical comparison, the small probability of considerably changing the momentum of a quark during the collision and the probability that the quarks may initially have a transverse component of momentum in the proper direction. The QCD theory should be able to calculate both of these small factors. They will depend on the energies of the protons and on the scattering angles, which need not be 90°. The recent polarization experiments indicate that leading-term predictions in perturbative QCD disagree with experiment.

Some theorists attempting to find new non-leading effects have investigated how in the exchange scenario the quarks in a proton acquire the large transverse component of momentum necessary to be going in the right direction before the collision. The leading term in the description of the structure of the proton has all three quarks moving in the same direction. There is a tiny probability that two quarks within the proton can have a single 'hard' collision with a high momentum transfer and come apart with high transverse momenta in opposite directions. But they can also have many 'soft' collisions with smaller momentum transfers which have a much higher probability and build up to high momenta from many successive small momentum transfers. The single hard scattering can be treated by perturbative QCD and conserves helicity. But many soft scatterings cannot be treated by perturbative QCD. Such 'non-perturbative' processes must be treated by non-leading terms and do not necessarily conserve helicity. Some model based on QCD is needed to estimate the importance of these effects, which might explain the single-helicity-flip transitions and other disagreements with perturbative QCD.

Thus, spin effects in high-energy hadron scattering are not trivial. They may shed light on how quarks are bound together by the forces of QCD to make hadrons. The problem of the description of spin effects in hadron scattering by QCD remains a challenge. □

Harry J Lipkin is at The Weizmann Institute of Science, Rehovot, Israel.

1. Court, G.R. *et al. Phys. Rev. Lett.* **57**, 507 (1986).
2. Gourlay, S.A. *et al. Phys. Rev. Lett.* **56**, 2244 (1986).
3. Crabb, D.G. *et al. Phys. Rev. Lett.* **41**, 1257 (1978).
4. Crosbie, E.A. *et al. Phys. Rev.* **D23**, 600 (1981).
5. Raymond, R.S. *et al.* in *Intersections Between Particle and Nuclear Physics* (ed. Mischke, R.E.) 1123 (AIP, Steamboat Springs, 1984).
6. Brodsky, S.J. & Lepage, G.P. *Phys. Rev.* **D23**, 2848 (1981).
7. Lipkin, J. in *School on Intermediate and High Energy Physics*, (ed. Falk, W.R.) (TRIUMF report, 1985).
8. Lipkin, H.J. *Ann. N.Y. Acad. Sci.* **452**, 79 (1985).
9. Federman, P., Rubinstein, H.R. & Talmi, I. *Phys. Lett.* **22**, 203 (1986).
10. Zeldovich, Ya. B. & Sakharov, A.D. *Yad. Fiz.* **4**, (1986).
11. Zeldovich, Ya. B. & Sakharov, A.D. *Sov. J. nucl. Phys.* **4**, 283 (1967).
12. Brodsky, S., Carlson, C. & Lipkin, H.J. *Phys. Rev.* **D20**, 2278 (1979).
13. Farra, G., Gottlieb, S., Sivers, D. & Thomas, G.H. *Phys. Rev.* **D20**, 202 (1979).
14. Barton, D.S. *et al. Brookhaven natn. Lab. Rep.* No. BNL-34771 (1984).

Immunology

Unravelling the control of B cells

Gerry G. B. Klaus

A CENTRAL feature of the humoral immune response is that when quiescent B lymphocytes encounter antigen they become activated, proliferate and differentiate into high-rate antibody-secreting cells. This process requires the participation of helper T lymphocytes and accessory cells such as macrophages. It has become increasingly clear that the helper function of T cells can be explained by their production of growth and differentiation factors that act on B cells. The precise functions of these molecules have, however, remained elusive until recently. Now, the complementary DNA encoding several of these lymphokines has been cloned and the sequences of two more are described in this issue[1,2], so the stage is set to unravel this complex process.

The existence of such factors was deduced about 15 years ago with the demonstration that culture supernatants from antigen or mitogen-stimulated mouse T cells contain factors (T-cell replacing factor(s) — TRF) which enable T-cell-depleted spleen cells to produce specific antibody responses[3]. Subsequently, such conditioned media revealed a bewildering variety of activities, such as activation of B cells from the resting state (B-cell stimulating factors — BSF); promotion of B-cell proliferation (BCGF); and/or differentiation to antibody secretion (BCDF). The development of monoclonal antibodies to individual lymphokines and of cloned T cells which secrete a limited array of factors has finally led to some progress in sorting out these activities. Within the past year, progress has markedly accelerated with the application of recombinant DNA technology.

This era was heralded by the cloning of the complementary DNA for murine BSF-1 (refs 4,5), followed quickly by that for the human homologue[6]. Murine BSF-1 has a relative molecular mass (M_r) of 20,000 (20K) and shows multiple activities; it acts on resting B cells and induces them to enter a transitional activation state characterized by the expression of high levels of antigens encoded by the class II major histocompatibility complex. In combination with antibodies to antigen receptors it stimulates these cells to proliferate. BSF-1 can also act as a rather selective B-cell differentiation factor by promoting the secretion of IgG1 and IgE antibodies by B cells pre-activated with lipopolysaccharide.

A second factor (BCGF-II, $M_r \sim 45K$) was initially recognized as a growth factor for a murine B-cell lymphoma (BCL$_1$)[7]. It has recently been shown to cause both DNA synthesis and antibody secretion (IgM and some IgG) in pre-activated normal mouse B cells[8]. Kinashi *et al.*[1] in this issue describe the cloning of the complementary DNA encoding murine BCGF-II, and show that the product also has TRF activity.

A third molecule is the M_r 20K BCDF, which induces antibody secretion in both Epstein–Barr virus-transformed human B cells and pre-activated normal cells, but has no BCGF activity[9]. The complementary DNA for this molecule has also now been cloned by Hirano *et al.*[2] but the murine homologue of this molecule appears not yet to have been identified. Interestingly, this human BCDF is the only one of these three molecules that so far seems to act only on B cells[9].

In contrast, murine BSF-1 is a growth factor for mast cells, and also for T cells, in a way distinct from interleukin-2 (ref. 10). The human homologue also has T-cell growth factor activity[4]. Murine BCGF-II also acts as an eosinophil differentiation factor (EDF, ref. 11), as does highly purified TRF (C. Sanderson, A. O'Garra and K. Takatsu, unpublished data), so there is little doubt that these three activities are produced by the same molecule. Lack of lineage specificity is common among cytokines. Interleukin-1 and interleukin-3 show diverse biological activities and interleukin-2 promotes B-cell proliferation and/or differentiation to antibody secretion (see ref. 12 for a review).

Not surprisingly, this rapid expansion of knowledge in B-cell factors has caused problems over nomenclature. BSF-1 and BCGF-II deserve to be added to the list of the three interleukins because of their multiple biological activities, and both were contenders for interleukin-4 (refs 4, 5, 11). As complementary DNA for BSF-1 was cloned first, this factor will retain that designation, and BCGF-II/TRF/EDF will presumably become interleukin-5 (as proposed by Kinashi *et al.*[1]). These designations, however, have yet to be officially ratified.

Does our current knowledge about these factors produce a coherent scheme of growth and differentiation control in antigen-stimulated B cells? The answer, unfortunately, is not yet. In murine B lymphocytes it is an attractive idea that resting cells become activated by the synergistic action of antigen plus BSF-1 produced by the cooperating T cell, and then become responsive to BCGF-II later in their cycle. Perhaps the balance between these two factors, in conjunction with unknown microenvironmental influences, will then determine which classes of antibody are secreted by the clonal progeny of a particular B cell, or if these progeny will instead revert to the resting state of memory B cells. It is significant that both BSF-1 and BCGF-II are growth and differentiation factors, suggesting that clonal expansion and maturation are coordinately regulated.

Unfortunately, what we know of human B cells does not fit this scheme. First, the genetic homologue of murine BSF-1 has not yet been shown to activate resting human B cells, but rather stimulates proliferation of pre-activated cells[6]. Second, it is likely that there is yet another type of B-cell growth factor for human cells (distinct from interleukin-2). Finally, the factor described by Kishimoto *et al.*[9] has no growth-promoting activity, suggesting compartmentalization of growth and differention in human B cells.

In conclusion, there still may be too many factors in the arena for comfort. Nevertheless, given the availability of the recombinant molecules, various questions can now readily be answered. For instance, are there functional B-cell subpopulations responsive to different factors? Do mixtures of individual lymphokines produce synergistic effects? The answers should not now take long to emerge. □

1. Kinashi, T. *et al. Nature* **324**, 70 (1986).
2. Hirano, T. *et al. Nature* **324**, 73 (1986).
3. Dutton, R.W. *et al. Prog. Immun.* **1**, 355 (1971).
4. Noma, Y. et al. *Nature* **319**, 640 (1986).
5. Lee, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 2061 (1986).
6. Yokota, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5894 (1986).
7. Swain, S.L. & Dutton, R.W. *J. exp. Med.* **156**, 1821 (1982).
8. O'Garra, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5228 (1986).
9. Hirano, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5490 (1984).
10. Mosmann, T.R. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 5654 (1984).
11. Sanderson, C.J., O'Garra, A., Warren, D.J. & Klaus, G.G.B. *Proc. natn. Acad. Sci. U.S.A.* **83**, 437 (1986).
12. Meidema, F. & Melief, C.J.M. *Immun. Today* **6**, 258 (1985).

Gerry G.B. Klaus is in the Division of Immunology, National Institute for Medical Research, London NW7 1AA, UK.

Neurobiology

Questions of taste and smell

Gordon M. Shepherd, Thomas V. Getchell and Charlotte M. Mistretta

RECEPTOR mechanisms in olfaction and taste have been tough nuts to crack for traditional methods of biochemistry and electrophysiology. The advent of molecular neurobiology and patch-clamp recordings is therefore most welcome, as was evident at a recent meeting*.

In olfaction, biochemical and molecular biological studies were made much easier by the development of isolated olfactory ciliary preparations containing enriched membrane fractions. Several molecular components localized in the fractions are believed to be associated with sensory transduction. The proposed sequence is initiated by binding of odour molecules (odogens) to membrane glycoproteins, such as gp95 (Chen, Z. & Lancet, D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1859; 1985), located in the cilia. The odogen–receptor complex is believed to activate the stimulatory GTP-binding protein (G_s) (D. Lancet, The Weizmann Institute). Although they appear to be less abundant, G_i and G_o binding proteins have also been identified in olfactory cilia using monospecific antisera (R.R.H. Anholt, Johns Hopkins University). Their functions may be related to receptor 'inhibition', or to neuromoldulatory effects of compounds such as substance P (J.F. Bouvet, Université Claude Bernard) or luteinsing hormone-releasing hormone (C.R. Wirsig and T.V.G., Wayne State University). G_s-deficient pseudohypoparathyroid patients have impaired olfactory perception compared with individuals with normal G_s activity (H.N. Wright, State University of New York).

Cyclic AMP is believed to be a second messenger that is activated during olfactory transduction and leads to excitation of receptor neurones. The activity of cyclic AMP in the olfacotry cilia preparations is 15 times that in the brain, and is increased substantially by GTP, its analogue GTPγS and forskolin, which indicates that the cyclase is activated by a G protein (P.B. Sklar, Johns Hopkins University). Certain molecules evoking different odour sensations also markedly increase cyclase activity. Others, which evoke putrid or solvent odours, do not stimulate this enzymatic activity. This observation suggests that parallel, non-adenylate cyclase-dependent transduction mechanisms are operable. For example, isolated olfactory ciliary preparations also exhibit increased phosphodiesterase activity when stimulated with L-alanine and GTP (R.C. Bruch and T. Huque, Monell Chemical Senses Center); this suggests that certain odogens stimulate phosphoinositide turnover by a G-protein-dependent mechanism.

A major question is to determine if cyclic AMP directly activates membrane channel proteins associated with the ion conductance changes responsible for the receptor potential, or if the effect is mediated by phosphorylation of specific proteins by phosphokinases in the cilia. Protein kinase C has been identified by phorbol ester binding and monospecific antisera (R.H. Anholt), suggesting that phosphorylation of the ion channels may be associated directly with conductance changes.

Patch-clamp recordings demonstrate that olfactory receptor neurones have interesting passive membrane properties that include high input resistances (2–6 GΩ) and low input capacitances 2.6–5 pF) (S. Firestein and F. Werblin, University of California, Berkeley; N. Suzuki, Hokkaido University). Membrane depolarizations are associated with a rapid inward Na^+ current followed by several outward K^+ currents (V.E. Dionne, University of California, San Diego; S. Firestein and F. Werblin). Near-threshold concentrations of odour molecules delivered by pressure injection elicit a 6-pA generator current that depolarizes the cell by 15 mV. There is a systematic relationship between the concentration of the odogen, the magnitude of the generator current and the number of action potentials initiated (S. Firestein and F. Werblin). These results support previous evidence that there is relatively close coupling between the conductance sites associated with ligand binding and those associated with impulse initiation (L. Masukawa, *J. Neurosci.* **5**, 128; 1985).

The relationship of the intermediate molecular stages of transduction and ion conductance changes has been investigated using G protein and cAMP analogues (D. Trotier and P. MacLeod, Laboratoire de Neurobiologie Sensorielle), which enhance conductive charges, supporting the biochemical evidence for an enzymatic cascade leading to the generator current. Single channel conductances in functionally reconstituted membrane proteins isolated from olfactory cilia incorporated in bilayer lipid membranes are enhanced by odour molecules and cyclic AMP (V. Vodyanoy and I. Vodyanoy, University of California, Irvine).

A protein of relative molecular mass 20,000 in the olfactory mucosa (J. Pevs-

---

**IX Symposium on Olfaction and Taste* and *8th Annual Meeting of the Association for Chemoreception Sciences* Snowmass Village, Colorado, 20–24 July 1986.

ner, Johns Hopkins University) binds several types of odour molecules with micromolar and millimolar affinities, is localized in olfactory glands and is released into the surface mucus when the glands are stimulated pharmacologically. The protein may facilitate diffusion of odour molecules through the mucus to receptor sites, or it may inactivate and clear odorants from the active sites. Odour molecules also stimulate secretion from olfactory glands by a cholinergic sensitive pathway, which may provide another mechanism by which odour molecules are cleared from olfactory receptor neurones (M.L. Getchell, Wayne State University).

Immunocytochemical techniques can help to elucidate the location of antigenic molecules in olfactory receptor neurones, which will help further studies on the coding of odour quality and the developmental sequences in neuronal maturation. A neurone-specific antigen (RB-8) in the rat has been characterized as a membrane-associated protein with a relative molecular mass of 125,000 (J.E. Schwob and D.I. Gottlieb, Washington University). In contrast to other regions of the nervous system, RB-8 stains the receptor neurone population nonhomogeneously: the axons of receptor neurones in the ventro-lateral region of the olfactory epithelium to the olfactory bulb are RB-8 positive, whereas axons from neurones with apparently identical morphology in the dorso-medial region of the olfactory epithelium are RB-8 negative. There is heterogeneity in cytochrome oxidase staining of olfactory receptor neurones during the development of normal and odour-deprived neonatal rat pups (P.E. Pedersen, Yale University). The further use of these two techniques should provide a clearer understanding of first principles of the functional organization of the peripheral olfactory system and its primary projections to the olfactory bulb, particularly as it relates to the coding of odour quality. Finally, monoclonal antibodies have been developed as specific probes to identify the emergence of membrane antigens during the development of olfactory receptor neurones (A.I. Farbman, Northwestern University; P.E. Pedersen). Two antibodies, Neu-5 and Neu-9, first bind on embryonic day 13 and stain the developing olfactory nerve bundles as well as the soma and dendrites. Another antibody, Neu-4, binds on embryonic day 14, the earliest time at which an odour-evoked response can be recorded from the olfactory mucosa. Further studies should lead to identification of specific membrane antigens that occur in associated with the development of odogen- and voltage-gated channels in receptor cell membranes.

In the taste field, there is particular interest in whether Na^+ channels are involved in transduction mechanisms for salt taste sensation, mainly deriving from studies where amiloride at least partially blocks the response to NaCl and LiCl but not to other salts.

Amiloride is now thought to block the sodium taste response completely, in contrast to previous reports of only partial blockade. Taste responses to sodium acetate and sodium bicarbonate are suppressed by 95–100 per cent after amiloride application on the tongue, whereas responses to NaCl and NaBr are suppressed by 65–90 per cent (B.K. Formaker and D.L. Hill, University of Toledo). Responses to NH_4Cl and ammonium acetate are not affected. Bretylium tosylate, an antifibrillary drug that opens amiloride-sensitive Na^+ channels, potentiates the taste of NaCl and LiCl, but has no effect on KCl or $CaCl_2$ when applied to the human tongue. The bretylium potentiation is inhibited by amiloride.

Other questions arise, such as why the monoclonal antibody 5B4 stains some, but not all, of the taste-bud cells in a bud. Could the antibody be binding a receptor protein that is present in cells only at certain stages of the turnover cycle? ☐

Gordon M. Shepherd is in the Section of Neuroanatomy, Yale University School of Medicine, New Haven, Conneticut 06510; Thomas V. Getchell is in the Depatment of Anatomy, Wayne State University, Detroit, Michigan 48021; Charlotte M. Mistretta is in the Department of Oral Biology, University of Michigan School of Dentistry, Ann Arbor, Michigan 48109, USA.

Geophysics

New radiocarbon dating system

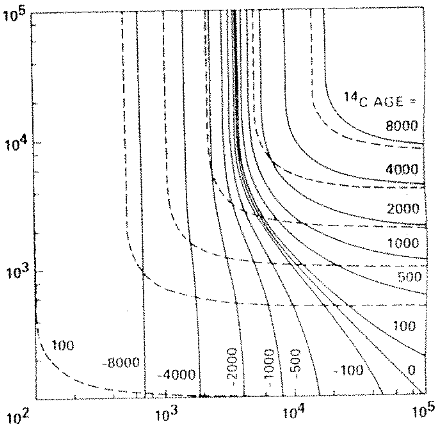
Roy Switsur

THE closest terrestrial analogue to martian conditions probably occurs in the arid valleys of southern Victoria Land, Antarctica; hypotheses suggesting that life exists on Mars are often based on this comparison. The discovery of cryptoendolithic microorganisms that survive in this inhospitable environment without completely adapting to its extremes add to this speculation. Recent studies of these microorganisms by McKay, Long and Friedmann[1] have led to a new approach to radiocarbon dating which is applicable to systems slowly exchanging carbon with the atmosphere, and which relates the true age to the apparent radiocarbon age and to the carbon-exchange rate.

Dry katabatic winds descend from the Antarctic ice plateau to create a lifeless desert where extensive areas of rock lie devoid of all snow or ice cover. The endolithic microorganisms colonize a narrow zone just beneath the surface of the porous granitic translucent rocks. Because of the hostile environment, photosynthesis can occur during only a few hours each year when conditions of moisture and light are suitable. A detailed analysis of this microbial ecosystem suggests that the organisms have very low rates of metabolism and may be very old[2].

The method of dating dead organic matter through its residual radiocarbon content, developed 40 years ago by W.F. Libby[3], is based on a closed system in which the carbon is unchanged through time apart from that lost by radioactive decay. The initial activity is produced through a chain of natural processes: a constant flux of primary cosmic rays impinging on the outer layers of the Earth's atmosphere generate secondary showers with high neutron content. These neutrons lose energy in elastic collisions as they descend until, at a height of about 12 km, they react with stable atmospheric ^{14}N nuclei to produce the radioactive ^{14}C isotope $^{14}N(n,p)^{14}C$.

Following rapid oxidation the radioactive CO_2 is mixed thoroughly, in a very short time compared with the half-life of the ^{14}C, with the existing CO_2 in the lower atmosphere. Most of the radiocarbon enters the oceans and becomes precipitated as a calcium carbonate ooze on the ocean floor. Part of the remainder becomes incorporated into plants through photosynthetic reactions. Of some 80 tons of the ^{14}C isotope that exists on the earth only approximately 3 tons remains in the biospheric reservoir, mainly in the form of photosynthetically produced cellulose. The carbon compounds thus produced by the green leaves of plants form the base of the food net for the animal life of the planet and their natural metabolism results in carbon exchange in equilibrium with the



True age in years (vertical scale) against mixing time-constant in years (horizontal scale). Solid line, with bomb effect; broken line, without bomb effect (from ref. 1.)

atmosphere. The ratio of the radioactive to stable carbon in these living organisms is 10^{-12}, producing a specific activity of about 6.8 picocuries per gram of carbon. On the death of the organism, carbon input ceases and the ratio decreases as the ^{14}C decays with a half-life of 5,730 years. Subsequent measurement of the radiocarbon concentration in the organic material permits the elapsed time since death to be calculated by use of the decay equation as long as there is no contamination by carbon from other sources.

It is well known that the assumption of the constancy of the atmospheric activity is erroneous. Measurements[4] of radiocarbon concentration in samples of growth rings of giant redwood trees of known age show perturbations of 1–2 per cent during the past 1,800 years, and further studies indicate that the natural variation is of the order of 10 per cent. It was realized that a calibration curve is essential for precise radiocarbon ages and much effort has been expended[5,6] in producing this from dendrochronologically dated Douglas firs and oaks.

Atmospheric activity is also affected by anthropological interventions. The cumulative effect of the combustion of fossil fuels (coal and oil) has diluted the activity to an extent estimated[7] to be about 3 per cent, so that wood grown, say, in 1950 would appear to be more than 200 years old. This effect is small compared with that of the intense production of radiocarbon by neutrons generated during the detonation of nuclear weapons in test programmes of the late 1950s and mid 1960s. During this short period atmospheric ^{14}C radioactivity increased by 100 per cent with a peak around 1963–64. The subsequent falling off in activity[8] is crucial to the model of McKay *et al.*

Although this model is applied specifically to the case of radiocarbon, it is capable of wider generalization to other radioisotopes that are diluted by mixing. The basic assumptions are that the carbon leaving the system has the same isotope composition as the total system; that the outward flux of carbon follows first-order reaction kinetics and hence is inversely proportional to the mixing time; and that the ^{14}C is a minute fraction of the total carbon. These assumptions allow a differential equation to be set up for the expression of ^{14}C continuity for a constant volume system and its integration. A further assumption, that the total carbon in the system remains steady over time so that the carbon flux in to the system balances that out, allows the fraction of modern carbon in the system to be estimated. In the absence of perturbations of atmospheric ^{14}C such as those described above, this estimate has a constant value of unity. The nuclear-bomb effect increases this atmospheric ^{14}C content so that the value rises above unity for the period subsequent to the commencement of bomb testing. McKay *et al.*[1] fitted a simple exponential curve to Nydal and Lövseth's monitoring data[8] for the post-1962 decrease in ^{14}C concentration to produce an equation for modern carbon content that does not depend on the true age of the system and may be evaluated for a given calendar date as a function only of the mixing time constant.

The authors also compare the analytical fit with a detailed numerical integration of the bomb radiocarbon measurements for a set of mixing times ranging from 10 to 10,000 years. The larger time-constants give the best agreement. Because the radiocarbon age computation implies a measurement of the ^{14}C content of the system and this is distorted by the temporal variation of modern carbon, the true age differs from the radiocarbon age. It is the value of the mixing time-constant that determines this difference. The figure shows a plot of true age versus mixing time-constant for several values of radiocarbon age, both with and without the bomb effect. The radiocarbon age lies on a locus of points bounded by the true age.

The problems to which the model might be applied require careful consideration but, apart from dating itself, it could possibly be of use for the study of the exchange rates of carbon in geological or biological systems. It might be used for modelling CO_2 in the atmosphere which requires a knowledge of the exchange rate of carbon between the reservoirs of soil and organic materials. Another intriguing possibility is the study of turn over time of carbon in individual living tissues such as bone collagen or brain. □

1. McKay, C.P., Long, A. & Friedmann, E.I. *J. geophys. Res.* **91,** 3836 (1986).
2. McKay, C.P. & Friedmann, E.I. *Antarct. J.U.S.* **19,** 170 (1984).
3. Libby, W.F. *Phys. Rev.* **69,** 671 (1946).
4. Willis, E.H., Tauber, H & Münnich, K.O. *Radiocarbon* **2,** 1 (1960).
5. Stuiver, M. *Radiocarton* **24,** 1 (1982).
6. Pearson, G.W., Pilcher, J.R., Baillie, MG. & Corbett, D.M. *Radiocarbon* (in the press).
7. Suess, H.E. *Science* **122,** 415 (1955).
8. Nydal, R. & Lövseth, K. *J. geophys. Res.* **88,** 3621 (1983).

Roy Switsur is in the sub-Department of Quaternary Research, University of Cambridge, Cambridge CB2 3RS, UK.

Archaeology

The environmentalist myth

Jared M. Diamond

MANY anthropologists and archaeologists believe that pre-industrial civilizations did far less damage to their environment than do their modern industrial counterparts. Derived from Rousseau's romanticized concept of the 'noble savage', this view holds that hunter-gatherers and neolithic agriculturists feel a reverence for nature, live in harmony with their environment, practise a conservation ethic and avoid the short-sighted destructive exploitation rampant in industrial societies. What reality does this postulated golden age of environmentalism possess? New evidence[1] adds support to the view that the golden age is a myth.

Recent archaeological discoveries have destroyed the assumption that Eden's hunters were conservationists. On every oceanic island for which we have adequate knowlege, the first arrival of humans was quickly followed by the extermination of all or most large animals, the best-known victims being the moas of New Zealand; the giant lemurs and elephant birds of Madagascar; and the flightless geese of Hawaii[2-4]. (It is still debated[2] whether the late Pleistocene extinctions of most large animals of North and South America and Australia were similarly caused by the arrival of humans.) More recently, it has become clear that pre-industrial societies were also prone to habitat destruction on a scale sufficient to destroy them.

One such case involves the collapse of the Polynesian society that erected the famous giant statues of Easter Island. This collapse proves to have been at least partly caused by deforestation, documented by pollen records[5,6]. When Polynesians reached the island around AD 400, they found it covered with palms, other trees and shrubs. By AD 1500 the human population had grown to about 7,000 and had

Barren grasslands and the famous monolithic statues of Easter Island. The scene was formerly a palm forest. (Courtesy of Jo Anne Van Tilburg and David Ochsner.)

erected some 245 stone platforms, carved more than 800 giant statues and destroyed the forest so completely that its tree species are now extinct[7]. Deforestation had three harmful consequences: soil erosion, hence lower crop yields; no trees to build canoes, hence lower yields of fish, the main protein source on Easter Island other than chickens; and inability to erect statues (weighing up to 85 tons), which had probably been raised to an upright position with the help of logs used as levers. The human population now exceeded the carrying capacity of the island. Warfare, cannibalism and slavery became chronic, a warrior class took over, spear-points were manufactured in enormous quantities and people reverted to living in caves for defence. By the time that Europeans discovered Easter Island, they found a panorama of destruction: the population may have crashed, many of the statues had been toppled and intentionally decapitated during battles between rival factions, and the island had been converted to barren treeless grassland.

A second example is the collapse of the Anasazi, one of the most advanced pre-Columbian civilizations in the United States. When European explorers reached the deserts of the American south-west, they found gigantic, uninhabited, multi-storied communal houses or pueblos. The dwellings in Chaco Canyon, for example, contained up to 5 stories and 650 rooms, were constructed with more than 200,000 roof beams weighing an average of 275 kg and were the largest and highest buildings in the United States until the advent of skyscrapers in the late nineteenth century. Archaeologists have shown that the dwellings were built about 1,000 years ago and abandoned in the twelfth century. The prevalent view, based on tree-ring evidence for a period of low summer rainfall, is that abandonment was forced by drought. But although drought may have contributed, the Anasazi themselves also inflicted fatal blows on the fragile desert environment.

Betancourt and Van Devender[8-11] used plant remains in radiocarbon-dated pack-rat middens to establish Chaco's vegetational history. When construction of the dwellings began in the early tenth century, the cliffs were covered with pinyon/juniper woodland, which was then cleared for firewood over a radius of 20 km during the next 200 years. The logs to construct the dwellings were mainly ponderosa pine, hauled in via roads from sites mostly more than 40 km away. As pines became depleted, the Anasazi began in AD 1030 to haul spruce, fir and Douglas fir logs from distances exceeding 75 km. In addition, Anasazi desert agriculture depended on gravity-flow irrigation without pumps, so that irrigation would have become impossible if channel erosion eventually lowered the water level below the field level. With the need to scour over a growing radius for construction timber and fuel, as well as the possible collapse of the irrigation system, the Chaco dwellings had to be abandoned. At Chaco Canyon, as on Easter Island, the environmental damage was irreversible.

Other cases deserve brief mention. It has long been debated whether overgrazing, removal of plant cover and soil erosion contributed to the decline of the ancient civilizations of Greece and the Middle East. A recent study[1] of the Peloponnisos suggests an answer in the affirmative by documenting at least four cycles of settlement, erosion and abandonment between 2000 BC and the Middle Ages. Abandonment was caused by slope erosion and valley siltation, resulting from shortened fallow cycles, clearing of steep slopes, domestic grazing animals and neglect of terracing and check dams. In Polynesia there was massive forest destruction in New Zealand[12] and Hawaii; degradation of forest to extensive fern savanna useless for agriculture in the Cook, Society, Marquesas and Mangaréva islands; and destruction of the xeric vegetation of Kahoolawe Island, causing a steep population decline and abandoment of all but a few coastal settlements after the sixteenth century[5]. Many other possible cases, such as the collapse of Mayan civilization and the desertification of the Sahara, cannot be evaluated without detailed evidence.

Although the lack of such evidence until recently surely helped the myth of an environmental golden age to persist, there may be another reason for the myth. It used to be common for industrial societies to treat pre-industrial peoples as savage brutes, meriting extermination or at least disposession of their lands. In reaction to this appalling history, others have idealized pre-industrial peoples as gentle conservationists. Both views are equally false. Ever since humans acquired the capacity to affect their environment, all societies have faced similar problems of managing fragile natural resources and have risked self-inflicted tragedies. ☐

1. Van Andel, T.H. *et al. Hesperia* **55**, 103 (1986).
2. Martin, P. S. & Klein, R. G. (eds) *Quaternary Extinctions* (University of Arizona Press, Tucson, 1984).
3. Diamond, J. M. *Nature News and Views* **298**, 787 (1982).
4. Steadman, D. W. & Olson, S. L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6191 (1985).
5. Kirch, P. V. *The Evolution of the Polynesian Chiefdoms* (Cambridge University Press, 1984).
6. Flenley, J. R. & King, S. M. *Nature* **307**, 47 (1984).
7. Van Tilburg, J. thesis, Univ. Calif. Los Angeles (1986).
8. Betancourt, J. L. & Van Devender, T. R. *Science* **214**, 656 (1981).
9. Samuels, N. L. & Betancourt, J. L. *Envir. Manag.* **6**, 505 (1982).
10. Betancourt, J. L. *et al.* in *Chaco Canyon Country, a Field Guide* (eds Wells, S. G. *et al.*) 207 (Geomorphological Field Group, 1983).
11. Betancourt, J. L. *et al. Am. Antiquity* **51**, 370 (1986).
12. McGlone, M. S. *Archaeol. Oceania* **18**, 11 (1983).

Jared M. Diamond is Professor of Physiology at the University of California School of Medicine, Los Angeles, California 90024, USA.

100 years ago

Observations on heredity in cats with an abnormal number of toes

THIS observation is concerned with a family of four male tabby kittens, all of which possessed the abnormality to a very market extent. This was the first family produced by a female tabby cat which, when born, was the most abnormal form which had come under my notice, possessing two extra toes on all the paws, i.e. seven on each fore-paw and six on each hind-paw. The extra toes (in the fore-feet) are those labelled A and B (in Figs. 1 and 2), and they confer the extraordinary breadth upon the foot. The most recently added is B, which is still partially coalesced with A, and has but one pad in common with it (Fig. 2)...There is seen to be an extra pad behind the additional toes, of which there is no trace in the normal foot." In the hind-paws (Figs. 5 and 6) "there is little doubt that the innermost toe I is the hallux lost in the normal foot...The second extra toe is that labelled A...On the under side (Fig. 6) all the toes have seperate pads, and there is an additional pad behind the extra toes," which is sometimes fused with that behind the normal toes.

From *Nature* **35**, 38; 11 November 1886.

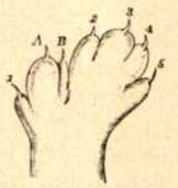


FIG. 1.—Right fore-paw from above, with extra toes.

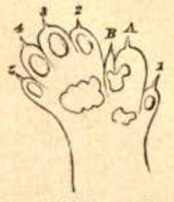


FIG. 2.—Right fore-paw from below, with extra toes.

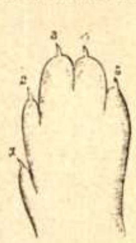


FIG. 3.—Right fore-paw from above, normal.

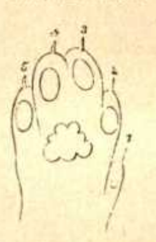


FIG. 4.—Right fore-paw from below, normal.

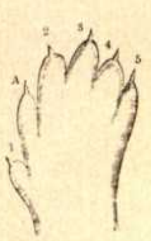


FIG. 5.—Right hind-paw from above, with extra toes.

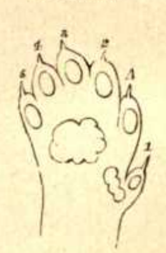


FIG. 6.—Right hind-paw from below, with extra toes.

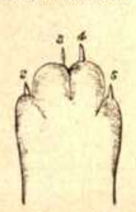


FIG. 7.—Right hind-paw from above, normal.

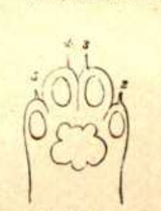


FIG. 8.—Right hind-paw from below, normal.

Where does the message begin?

SIR—The remarkable accomplishment of cloning and sequencing the gene responsible for the X-linked form of the phagocytic disorder chronic granulomatous disease (X-CGD) on the basis of its chromosomal location alone, without reference to any protein, has recently been reported in these pages by Royer–Pokora *et al.*[1]. With the amino acid sequence predicted from the gene sequence, it should now be possible to identify and characterize the previously unknown protein responsible for this disorder. However, it seems possible that the true amino terminus of the protein has not been identified.

Royer–Pokora *et al.* conclude that translation of the X-CGD messenger RNA probably initiates at the second in-frame AUG codon, at position 322, because it occurs in a sequence that more closely resembles the consensus sequence for functional initiator codons, defined by Kozak[2] on the basis of observed sequences, than does that at the first AUG at position 208. The primary difference between the two is that the AUG at position 322 has an A nucleotide three positions upstream, as do 79 per cent of functional initiation codons, whereas at the first AUG there is a U, which is hardly ever observed at this position[2]. On the other hand, 95 per cent of known messenger RNA molecules have translation initiated at the most proximal AUG codon[2], and the importance of the most proximal position has been demon-strated directly[3]. The basis of the nucleotide preference observed three nucleotides upstream is not known. Therefore, the choice of neither initiation codon is compelling from just these considerations.

The reason for suggesting that translation of the X-CGD messenger RNA initiates at the first AUG is that the 38 extra amino acid residues indicated by the messenger RNA sequence are not random. In particular, the first 20 residues closely resemble the signal peptides that have been implicated in the co-translational passage of proteins into the endoplasmic reticulum as the first step in targeting the protein for secretion from the cell, insertion into membranes, or sequestration into at least some cytoplasmic compartments[4-6]. The pertinent properties are that this sequence;

Met-Leu-Ile-Leu-Leu-Pro-Val-Cys-Arg-Asn-Leu-Leu-Ser-Phe-Leu-Arg-Gly-Ser-Ser-Ala-

is positively-charged, has predominantly hydrophobic amino acids from residues 2 to 15, and ends with five polar residues typical of signal peptide cleavage sites. The only unfavourable aspects of the sequence are the Arg and Asn residues at positions 9 and 10, respectively, within the hydrophobic segment, but such residues are observed at low frequencies in similar positions in known signal peptides[7]. The predictive scheme of Von Heijne[7] considers the frequencies with which the 12 amino acids proximal to the signal peptide cleavage site, plus the two distal, are observed in known signal sequences. The relative likelihood of the first 20 residues constituting a signal peptide, with cleavage after Ala 20, was calculated to be + 5.6, which is well within the range for known signal peptides and outside that found for the terminal amino-acid sequences of cytoplasmic proteins[7]. The first eight residues are not considered in this calculation but are very like those in known signal peptides.

The nature of the X-CGD disorder is consistent with the protein responsible being situated in granules of phagocytes[8]. Proteins designated for such cell sites are often synthesized with signal sequences and are directed to the appropriate cell compartment after entering the endoplasmic reticulum[9-11].

THOMAS E. CREIGHTON
Medical Research Council,
Laboratory of Molecular Biology,
Cambridge, CB2 2QH, UK

1. Royer–Pokora, B. *et al. Nature* **322**, 32–38 (1986).
2. Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
3. Sherman, F., Stewart, S.W. & Schweingruber, A.M. *Cell* **20**, 215–222 (1980).
4. Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1496–1500 (1980).
5. Perlman, D. & Halvorson, H.O. *J. molec. Biol.* **167**, 391–409 (1983).
6. Von Heijne, G. *J. molec. Biol.* **184**, 99–105 (1985).
7. Von Heijne, G. *Nucleic Acids Res.* **14**, 4683–4690 (1986).
8. Tauber, A.I., Borregaard, N., Simons, E. & Wright, J. *Medicine* **62**, 286–309 (1983).
9. Faust, P.L., Kornfeld, S. & Chirgwin, J.M. *Proc. natn. Acad. Sci. U.S.A* **82**, 4910–4914 (1985).
10. Shibahara, S. *et al. Nucleic Acids Res.* **14**, 2413–2417 (1986).
11. Lobe, C. *et al. Science* **232**, 858–861 (1986).

SIR—We appreciate the comments and suggestions from Creighton[1] and from Ashworth *et al.*[2] regarding our predictions for the protein product encoded by the gene responsible for X-linked chronic granulomatous disease[2].

Creighton suggests that a non-consensus ATG at position 208 of our cDNA sequence might be the correct translation initiation site (rather than the consensus ATG at position 322 we favoured) and that the additional N-terminal extension encodes a signal peptide based on a predictive scheme of von Heijne[4]. As stated in our paper, no certain choice between these ATGs can be made from cDNA sequence alone. Our analyses of the sequence, albeit without the recent approach of von Heijne, led us to favour the ATG at 322. Creighton's proposal needs to be very carefully considered, but in light of our recent finding (R. Brown and S.H.O., unpublished data) that an intron interrupts the alanine codon he suggests as a signal peptide cleavage site.

Whatever the final resolution, the uncertainty in ATG assignment does not materially affect experimental approaches to identification of the protein *in vitro*. For one, the size difference between the proteins predicted from initiation at either ATG (minus the signal peptide if ATG 208 were used) is small. In addition, we are obliged to examine all cellular fractions to localize the protein to a specific compartment. Finally, initial results with antisera directed against a portion of the predicted amino acid sequences identify an *in vivo* protein located in membrane-rich fractions of granulocytes (M. Dinauer and S.H.O. unpublished data). Expression of transferred cDNAs will permit definitive assignment of the initiator codon and the possible role of the signal peptide proposed by Creighton in the targeting of the protein within the cell.

Although the suggestion of Ashworth *et al.*[2] that a region near the C-terminus of the predicted protein resembles a haem-binding site of a cytochrome P450 is intriguing, the significance of the limited resemblance they note is unclear.

STUART H. ORKIN
Division of Hematology, Children's Hospital,
Howard Hughes Medical Institute,
Harvard Medical School,
Boston, Massachusetts 02115 USA

1. Creighton, T. *Nature* **324**, 21 (1980).
2. Ashworth, A., Shephard, E.A.& Phillips, I.R. *Nature* **322**. 599 (1986).
3. Royer Pokora, B. *et al. Nature* **322**, 32–38 (1986).
4. Von Heijne, G. *Nucleic Acids Res.* **14**, 4683–4690 (1986).

Occam and mankind's genetic bottleneck

P442

SIR—In defending their support for and amplification of the view[1] that non-African human populations are descended from African populations via an emigrant stock that was of very small size, Jones and Rouhani[2] have exposed the weakness of their argument. It is true that such a bottleneck is an inevitable implication of their simple model. However, that model assumes that the β-globin non-coding sequences of DNA on which their argument is based are unaffected by selection, either directly or by linkage ('hitch-hiking'). This assumption is defended on two grounds. The first is that neutrality of non-coding sequences is implicit in most hypotheses of molecular evolution. But there is actually no evidence for this fashionable assumption. It is adopted merely because of the prejudices of those advancing the hypotheses, because it allows phylogenetic conclusions to be reached in molecular biology laboratories rather than at palaeontologists' benches, and possibly because it makes the models mathematically easier.

Jones and Rouhani tell us that their second reason for assuming neutrality is that, since there is no information on any selective forces affecting the variants they discuss, they "have, with Occam, chosen the more straightforward alternative"[2]. Let us not argue fruitlessly about what William of Occam really meant. In practice, science progresses by attempting to explain phenomena simply, by trying to simplify further, and by adding complexity only when necessary. But what is simpler in this case — assuming no selection and a bottleneck, or assuming selection and no bottleneck? Conceptually the two hypotheses are equally simple: the maths can be simplified either by leaving out selection or by assuming the population size to be effectively infinite. Indeed, the available data are consistent with innumerable combinations of values of the relevant variables — population number, selective coefficients, time and original allele-frequencies. One can, of course, reach conclusions about the values of some of them by making arbitrary assumptions about the values of others but there is no justification for doing so. We just have to admit our ignorance.

Jones and Rouhani are prepared to invoke selection when they need it: they argue against parallel evolution in various populations by recalling Sewall Wright's calculations of the unlikelihood of a chromosome rearrangement becoming established in a population if it is at a selective disadvantage in the heterozygous condition. Do they consider the β-globin sequences to be neutral or do they not?

JEREMY J.D. GREENWOOD
Department of Biological Sciences,
The University of Dundee,
Dundee, DD1 4HN, UK

1. Wainscot, J.S. *et al. Nature* **319**, 491-493 (1986).
2. Jones, J.S. & Rouhani, S. *Nature* **322**, 599-600 (1986).

Chernobyl fallout in size-fractionated aerosol

SIR—The Chernobyl accident first produced elevated radioactivity in ambient air in Switzerland on 29 April. Daily high-volume samples showed that the highest activities occurred on 1 May and had already decreased by two orders of magnitude by 8 May. The maximum values observed of ^{137}Cs (2 Bq m^{-3}) are about half of those reported from Studsvik[1].

To understand the interaction of these airborne fission nuclides with the environment, we need to know the carrier responsible for their long-range transport. We therefore measured activity distributions on size fractionated aerosol samples for typical fission products such as ^{131}I, ^{103}Ru, ^{132}Te and ^{137}Cs, using two different commercially available cascade impactors.

From 30 April to 13 May, an Andersen sampler collected aerosols in Spiez (total 570 m^3) and, from 2 May to 8 May, a low pressure Berner impactor[2] was intermittently in operation in Zurich (total 170 m^3). The impactors fractionate airborne particles into different size classes.

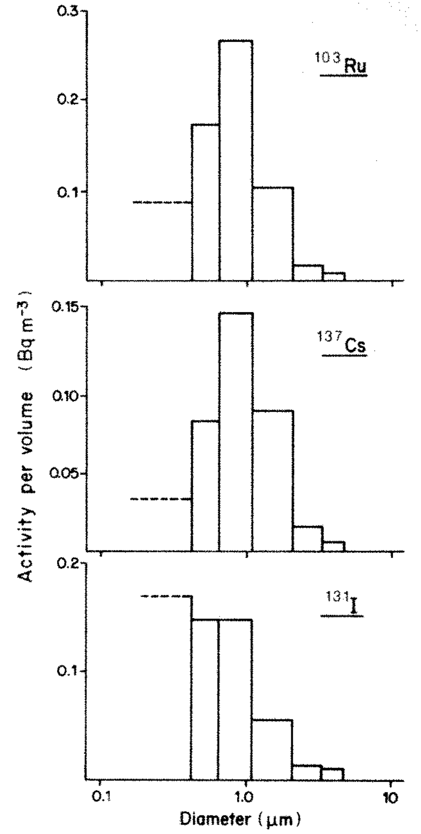


Fig. 1 Activities of ^{103}Ru, ^{137}Cs, and ^{131}I in size fractionated aerosol samples in Bq m^{-3} sampled at Spiez from 09:00 h April 30 to 09.00h May 13 using an Andersen sampler. Note that the low activities are due to averaging over the 14 day sampling period.

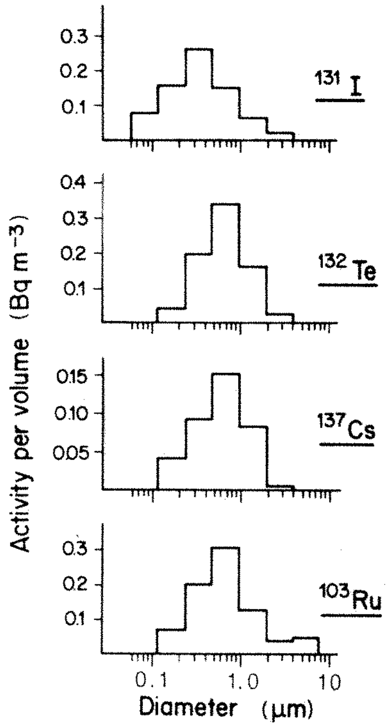


Fig. 2 Activities of ^{131}I, ^{132}Te, ^{137}Cs, and ^{103}Ru in size-fractionated aerosol samples in Bq m^{-3} sampled at Zurich-Hoenggerberg from 09:15 h—May 23:15 h May 6th using a Berner cascade impactor.

As Fig. 1 shows, ^{131}I has a rather different distribution from the other nuclides. Most ^{131}I is found in the size fraction on the backup filter (<0.47 μm), whereas ^{103}Ru and ^{137}Cs show a pronounced maximum at 0.93 μm (geometric mean diameter). The Berner impactor (Fig. 2). gives a similar distribution with the maximum for iodine at 0.35 μm and at 0.71 μm for ^{132}Te, ^{137}Cs and ^{103}Ru. Different distributions are due to the different cutoff diameters of the devices. The distributions for ^{137}Cs, ^{132}Te and ^{103}Ru are very similar to the concentration patterns of prominent ions such as NH_4^+ or SO_4^{2-} and NO_3^- as determined by ion chromatography (data not shown).

Measurements in Sweden [1], Finland[3], and Germany [4] have shown that ^{131}I was mainly (~ 80%) present in the gas phase. The adsorption of gaseous species onto aerosols is strongly surface correlated[5,6]; measured surface distributions for environmental aerosols usually peak at size fractions with particle diameters ≤0.4 μm (refs 7,8). We therefore conclude that the size distribution observed for ^{131}I does not reflect the primary particles released at Chernobyl but rather the surface distribution of the aerosols at the sampling site.

By contrast, ^{137}Cs, ^{131}Te and ^{103}Ru were presumably ejected as particles or attached very early to aerosols and grew by coagulation with other particles during transport. It is therefore reasonable that their size distribution resembles that of the main inorganic ions (sulphate, nitrate and ammonium) which may also be transported over long distances[9].

In contrast to the findings of Bondietti and Brantley[10], we found no increase of the activity median aerodynamic diameter during the measurement period. We think that this increase may not reflect the original release at Chernobyl, but be due to transport effects. Note that the size distribution of ^{137}Cs, for example, from the Chernobyl fallout is very similar to the size distribution in the fallout of nuclear weapons tests[11], in contrast to the size distribution found in the stratosphere, which is shifted towards smaller particles[12].

We have also measured fission products in sequential samples of the first rain at Zurich after 29 April. The activity distributions of the radionuclides were very similar to the concentration distributions of the ions, from which we conclude that the depletion mechanisms are quite similar for both groups. The activity ratio of ^{131}I to the other radioisotopes, such as ^{131}I/^{103}Ru in rain is about 5, whereas in the impactor samples it is near to one (integrated values from Fig. 1 or 2). This observation also indicates that in the atmosphere ^{131}I was mainly present as gaseous species and is depleted with high efficiency by rain.

D.T. JOST
H.W. GÄGGELER
U. BALTENSPERGER
Swiss Federal Institute for Reactor Research
CH-5303 Wuerenlingen, Switzerland
B. ZINDER
Atmospheric Physics ETH
CH-8093 Zurich, Switzerland
P. HALLER
AC-Zentrum
CH-3700 Spiez, Switzerland

1. Devell, L. *et al. Nature* **321**, 192–193 (1986).
2. Berner, A. & Luerzer, C. *J. Phys. Chem.* **84**, 2079–2083 (1980).
3. Finnish Centre for Radiation and Nuclear Safety *et al. Second interim report: Radiation situation in Finland from 5 to 16 May 1986* (STUK-B-VALO 45, Helsinki 1986).
4. *Umweltradioaktivitaet und Strahlenexposition in Suedbayern durch den Tschernobyl-Unfall* (GSF-Bericht 16/86) (Gesellschaft fuer Strahlen- und Umweltforschung, Munich, 1986).
5. Raunemaa, T. & Hautojärvi, A. *J. Aerosol Sci.* **15**, 123–132 (1984).
6. Kruger, J. & Nöthling, J.F. *J. Aerosol Sci.* **10**, 571–579 (1979).
7. Seinfeld, J.H. *Atmospheric Chemistry and Physics of Air Pollution*, 289 (Wiley, New York, 1986).
8. Whitby, K.T. *et al. J. Colloid Interface Sci.* **39**, 177–204 (1972).
9. Weiss, R.E. *et al. Science* **195**, 979–981 (1977).
10. Bondietti, E.A. & Brantley, J.N. *Nature* **322**, 313–314 (1986).
11. Lockhart, L.B., Jr. *et al. J. geophys. Res.* **70**, 6033–6041 (1965).
12. Persson G. & Sisefsky, *J. Hlth Phys.* **21**, 421-428 (1971).

Alimentary restrictions and ^{131}I in human thyroids

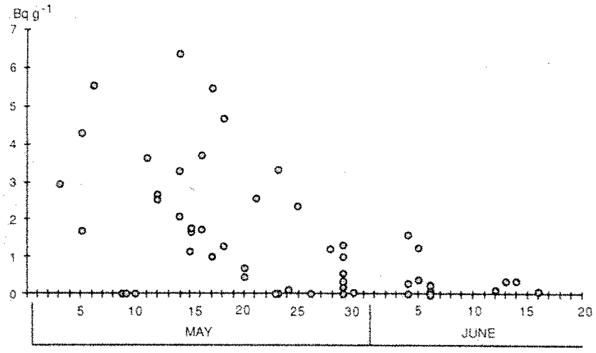


Fig. 1 ^{131}I activity as measured in 51 thyroid specimens in the period following the Chernobyl accident.

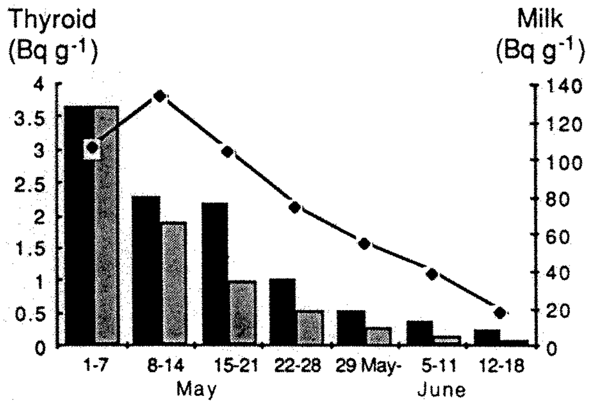


Fig. 2 Mean weekly ^{131}I activity in milk (continuous line) and in thyroid (black columns) during the 7 weeks following arrival of the radioactive plume. Stipled columns indicate the theoretically expected decay of radioactivity, starting from the levels detected in thyroid during the first week of monitoring.

SIR—Hill *et al.*[1] have reported measurements of ^{131}I during the period following the Chernobyl reactor accident. We have directly determined ^{131}I activity in specimens of thyroid gland obtained at autopsy from 51 adult individuals dying of accidental causes in the area of Genoa. Sampling started on 3 May 1986, 2 days after arrival of the radioactive plume over the area and one day before an abundant rainfall.

Each thyroid specimen was placed in the central well of a shielded NaI(T1) detector connected with a 1,024-channel spectrometer; radioactivity was measured against a ^{131}I standard having the same geometry.

Differences between individuals (Fig. 1) presumably reflect variable exposure. High levels of ^{131}I were detected from the beginning of radioactive fall-out. The decay of radioactivity in the following weeks was less pronounced than expected on the basis of the effective half-life of ^{131}I in thyroid (7.55 days) (Fig. 2), inferred from its biological (120 days) and physical (8.05 days) half-life.

Although these data are consistent with some persistence of the exposure to ^{131}I, the progressive decrease of radioactivity in the thyroid contrasts with the initial increase of ^{131}I in milk samples from the same area (Fig. 2), presumably reflecting the efficacy of alimentary restrictions introduced in Italy on 3 May. It is also significant that, on withdrawal of restrictions (16 May), the trend towards the loss of radioactivity in the thyroid was temporarily discontinued (Fig. 2).

The mean absorbed β dose was inferred from the ^{131}I activity measured in thyroid specimens, using an *ad hoc* conversion factor[2]. During the first 7 weeks following arrival of the radioactive plume, the mean (± SD) weekly absorbed doses were 3.66 ± 1.68, 2.31 ± 2.12, 2.20 ± 1.61, 1.02 ± 1.37, 0.55 ± 0.58, 0.38 ± 0.51, and 0.24 ± 0.17 μSv g^{-1} respectively. Assuming an average weight of 20 g per thyroid, it can be estimated that the cumulative load absorbed during that period accounted for a mean value of 207 μSv per individual. These estimates are higher than those made for adult individuals in southern United Kingdom, where the environmental contamination, as deduced from milk ^{131}I activity, was substantially lower than in our region[1].

P. ORLANDO
G. GALLELLI
F. PERDELLI
S. DE FLORA
Institute of Hygiene
R. MALCONTENTI
Institute of Forensic Medicine,
School of Medicine,
University of Genoa,
16132 Genoa, Italy

1. Hill, C.R., Adam, I., Anderson, W., Ott, R.J. & Sowby, F.D. *Nature* **321**, 655-656 (1986).
2. Snyder, W.S., Ford, M.R., Warner, G.G. & Watson, S.B. *MIRD pamphlet No. 11* (Society of Nuclear Medicine, New York, 1969).

Radioactive fallout swipe samples from Chernobyl

SIR—Significant quantities of radioactive fallout from Chernobyl were detected in aircraft swipe samples throughout the month of May, initially on flights from Moscow and, later, from London, New York and Tokyo. Swipe samples have provided valuable information on atmospheric nuclear weapon tests[1,2] in the past. The fallout from Chernobyl was measured in the swipe samples collected at Bombay from 2 May onwards. Though the radioactive plume rise at the accident site has been estimated[3] to be around a kilometre, atmospheric dispersion with time could carry the activity to higher altitudes which enabled its detection.

In all, 20 isotopes; ^{95}Zr, ^{95}Nb, ^{99}Mo, ^{103}Ru, ^{106}Ru, ^{110m}Ag, ^{111}Ag, ^{125}Sb, ^{127}Sb, ^{132}Te, ^{131}I, ^{134}Cs, ^{136}Cs, ^{137}Cs, ^{140}Ba, ^{140}La, ^{141}Ce, ^{144}Ce, ^{147}Nd and ^{239}Np were measured by high resolution gamma spectroscopy. Although the nature of the sampling does not permit quantification with respect to volume of air, useful information could be obtained by studying suitable radioisotope ratios.

Our results are generally similar to those reported by others[3-6]. Different isotopes of caesium, iodine, ^{132}Te, ^{140}Ba, ^{140}La, ^{103}Ru and ^{99}Mo were the major contributors to the activities measured in the samples. Comparisons with similar swipe samples collected after the last atmospheric test in 1980, which was a low altitude H-bomb explosion, show that in the present case the ^{137}Cs levels are about an order of magnitude higher and ^{103}Ru an order of magnitude lower. The levels of ^{132}Te and ^{99}Mo, are 10–100 times lower. The other major difference is that the ^{95}Zr concentration is 1000 times less, than from the samples after the last nuclear test.

Table 1 Activity Ratios (on 26.4.86) of fission products and ^{239}Np to ^{137}Cs in 5 swipe samples

	Origin of flight and the date of collection at Bombay				
Ratio of element to ^{137}Cs	Moscow 2.5.86	New York 5.5.86	New York 11.5.86	Tokyo 11.5.86	Tokyo 29.5.86
^{95}Zr	0.022	—	0.002	0.003	0.002
^{95}Nb	0.05	0.04	0.025	0.026	0.019
^{99}Mo	0.41	1.04	1.48	1.81	—
^{103}Ru	0.19	0.95	1.09	1.72	1.27
^{106}Ru	0.065	0.18	0.24	0.35	0.29
^{110m}Ag	0.008	0.013	0.014	0.016	0.014
^{111}Ag	0.8	—	—	—	—
^{125}Sb	0.017	0.04	0.03	0.05	0.03
^{127}Sb	0.24	0.43	0.51	0.56	—
^{132}Te	4.05	10.2	11.5	18.4	13.2
^{131}I	1.84	1.06	0.65	2.13	1.13
^{134}Cs	0.50	0.47	0.51	0.56	0.50
^{136}Cs	0.19	0.17	0.20	0.22	0.20
^{140}Ba	0.40	0.69	0.62	0.90	0.63
^{140}La	0.50	0.79	0.87	1.19	0.82
^{141}Ce	0.04	0.02	0.04	0.05	0.05
^{144}Ce	0.07	—	0.02	0.02	0.03
^{147}Nd	0.014	—	—	—	—
^{239}Np	0.42	0.48	4.9	7.1	—
^{137}Cs activity (becquerels)	116	152	315	1241	427

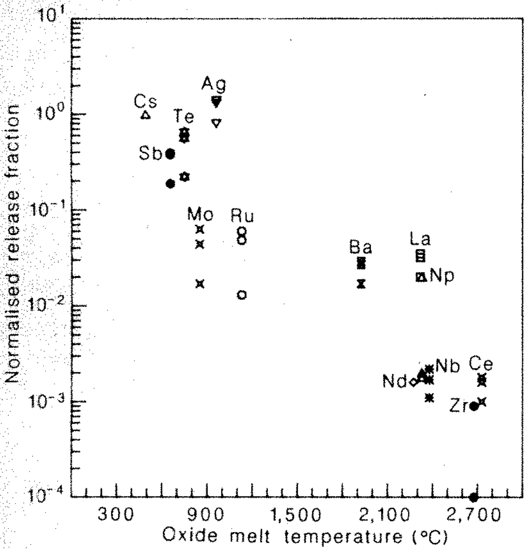


Fig. 1 Activity ratios of selected fission products with respect to ^{137}Cs. The activity of ^{137}Cs in the samples are also shown.

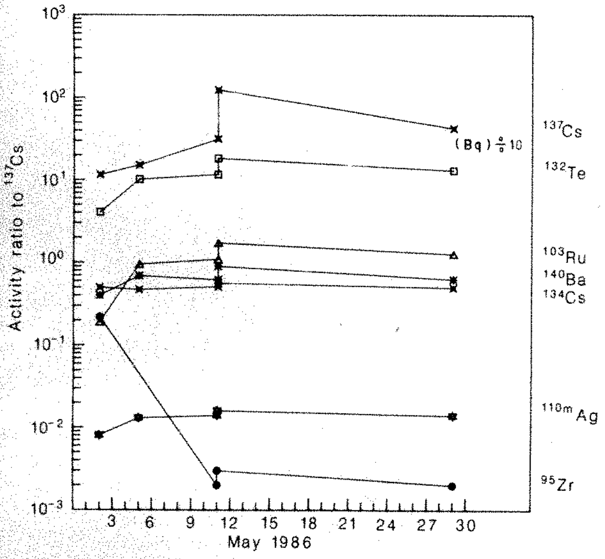


Fig. 2 Release fraction with respect to Cs plotted against fission product oxide's melting point.

The isotope activity ratios to ^{137}Cs (as on 26 April) in 5 of the samples collected during May are given in Table 1 and a few selected ratios are plotted in Fig.1. The ratios of elements whose oxides have high melting temperatures (for example Zr, Nb and Ce) with respect to ^{137}Cs were higher in the first sample collected on 2 May, probably representing the initial release from the reactor. Subsequent samples show higher relative concentrations of Ru, Te, Mo and Sb and, to a lesser extent, of Ba, La and Ag. The ^{131}I/^{137}Cs ratio variations in these samples did not show any set pattern, perhaps because much of the ^{131}I may be in vapour form[4].

Two estimates of the reactor fuel burn up have been reported[4,6], based on ^{134}Cs/^{137}Cs and ^{103}Ru/^{106}Ru ratio values of 0.5 and 4.5 respectively; we observed similar values for the ratios. We also obtained other ratios from the isotope pairs of the same elements such as ^{127}Sb/^{125}Sb, ^{136}Cs/^{134}Cs and ^{141}Ce/^{144}Ce, the last of which varied somewhat from sample to sample. We obtained reactor core inventories from the ORIGEN 2 code[7] for 1.8% enriched uranium Candu type reactor and compared them with the observed activity ratios. They correspond to a fuel exposure of around 650 d (700 d if 2.0% enriched U was used) though our ^{134}Cs/^{137}Cs also gave fuel exposure of 400d.

The release fractions, normalized to ^{137}Cs, vary with the melting point of the oxides of the fission products for the first three samples. The fractions for Sb and Te agree well with those documented for the PWR reactor core melt accident[8] while for other elements marked differences are observed. Deviations start from Mo onwards and increase with increasing oxide melting point, suggesting that there may have been only a partial core melt (Fig.2).

We thank V.K. Sundaram, Health Physics Division, for the ORIGEN 2 results and Dr V.K. Shukla, G.N. Shaikh and T.N. Mahadevan for help in sample collection.

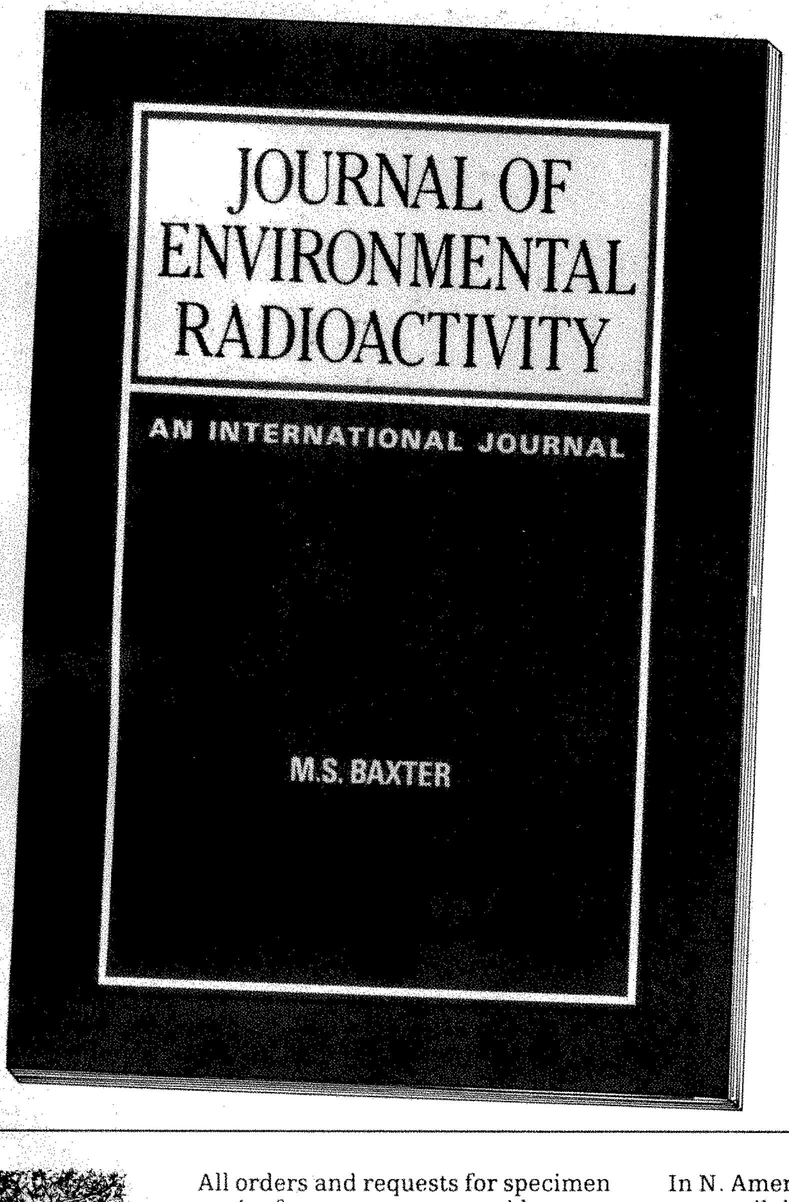
S. SADASIVAN
U.C. MISHRA

Air Monitoring Section,
Bhabha Atomic Research Centre,
Bombay 400 085, India

1. Mishra *et al. J. sci. indus. Res.* **33,** 216–221 (1974).
2. Sadasivan, S. & Mishra, U.C. *J. geophys. Res.* **87,** 7343–7345 (1982).
3. *Interim report on radiation situation in Finland from April 26 to May 4,* (STUK-B-VALO 44) (Finnish Centre for Radiation and Nuclear Safety, Helsinki, 1986)
4. Devell *et al. Nature* **321,** 192–193 (1986).
5. Fry, F.A., Clarke, R.H. & O'Riordan, M.C. *Nature* **321,** 193–195 (1986).
6. Hohenemser *et al. Nature* **321,** 817 (1986).
7. Croff, A.G. *Nucl. Tech.* **62,** 335–352 (1983).
8. *U.S. Nuclear Regulatory Commission Reactor Safety Study* (WASH 1400, NUREG 75/014) (Washington, 1975).

For research papers on Chernobyl

JOURNAL OF ENVIRONMENTAL RADIOACTIVITY



Editor:
M.S. Baxter, *Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, Scotland*

"Although many periodicals will take articles on environmental aspects of radioactivity, there is a good case for a journal specifically concerned with the subject. Now we have one, Journal of Environmental Radioactivity. Most of the papers published to date are naturally concerned with radioactive pollution from materials arising from nuclear bomb tests or from nuclear industry."

John Fremlin
Nature, September 1985

This review was, of course, written before the incident at Chernobyl. Many journals have already published "immediate" papers on its effects. However, there are and will be "long-term effect" papers which will be published continuously over the next few years.

Journal of Environmental Radioactivity is the obvious medium for these papers which will be featured regularly.

The journal accepts original contributions of 2,000 to 4,500 words (plus tables and figures); review articles, notes and conference reports are also accepted. There are no page charges.

For further information write to the publisher or the Editor.

Subscription Information
Frequency: 6 issues in 1 volume.
Volume 5 1987

Price: U.K. delivery £95.00.
Overseas delivery £105.00/$174.00
The £ sterling price is definitive.
US$ prices are subject to exchange rate fluctuations.

All subscriptions are airspeeded at no extra cost.

ISSN: 0265-931X

All orders and requests for specimen copies from customers outside N. America:
Elsevier Applied Science Publishers
Crown House, Linton Road,
Barking, Essex, IG11 8JU, UK.

In N. America free specimen copies are available from:
Journal Information Centre,
Elsevier Science Publishing Co. Inc.,
52 Vanderbilt Avenue, New York,
NY 10017, USA.
Telephone: (212) 916 1250

ELSEVIER APPLIED SCIENCE PUBLISHERS

ANNOUNCING TWO NEW JOURNALS FOR 1987

Journal of Chemical Neuroanatomy

Editor: H.W.M. STEINBUSCH

Laboratorium Voor Farmacologie,
Vrije Universiteit, Amsterdam, The Netherlands

Advances in investigative techniques and methods have led to vastly-improved microscopic resolution of the structure and function of neural cells, and especially information about receptor and transmitter localisations and their mechanisms of action. This major new publication in the neurosciences will specialize in original research papers specifically relating to transmitter-characterized systems in the central and peripheral nervous tissue, with a special emphasis on methodological papers. Inter-disciplinary, developmental, evolutionary, ageing and behavioural investigations relating to one or more neuronal systems will also be encouraged.

To be published bi-monthly. Prices to be announced.

Microbial Ecology in Health and Disease

Editors-in-Chief

P. BORRIELLO
Division of Communicable Diseases
Clinical Research Centre, Harrow UK

C. RUSSEL
Department of Oral Medicine
University of Manchester, UK

The journal will consider manuscripts primarily on the human microbial flora, but also that of other animals, and its developments and interactions.

Topics will include

★ investigative methods ★ animal and *in vitro* models ★ the effect of anti-microbials or diet on the commensal flora or its development ★ the role of immunological and other mechanisms which help maintain a stable flora ★ the clinical applications of the commensal flora in the treatment and prevention of disease ★ alterations in the host environment (eg. colostomy, dental treatment).

To be published quarterly. Prices to be announced.

Please write for more information to

Dept JSL, John Wiley & Sons Ltd

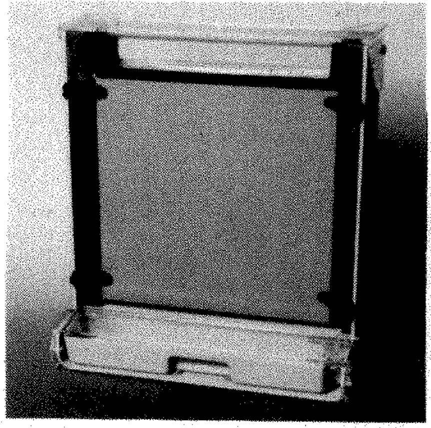
Baffins Lane, Chichester, Sussex PO19 1UD Telephone (0243) 784531 Telex 86290

Wiley Life Sciences · Wiley Life Sciences · Wiley Life Sciences · Wiley Life Sciences · Wiley Life Sciences · Wiley Life Sciences · Wiley Life Sciences ·

Reader Service No.9

Put a smile on your face... ...not on your Gel

The Model S2 Sequencing apparatus from Bethesda Research Laboratories combines safety and performance in a simple, easy to use system. Now achieve clear, readable 'no smile' autoradiograms with ease



The S2 incorporates

- an aluminium plate for even heat distribution
- drainage from the upper buffer chamber to the lower
- sharkstooth combs
- safety interlock buffer chamber lids
- unique plate fasteners for easy assembly
- removable electrode

GIBCO BRL
Life Technologies

United Kingdom GIBCO Limited, PO Box 35, Trident House, Renfrew Road, PAISLEY PA3 4EF. Renfrewshire, Scotland. Telephone: 041-889 6100 Telex: 777968/777029
United Kingdom GIBCO Limited, Unit 4, Cowley Mill Trading Estate, Longbridge Way, UXBIRDGE UB8 2YG, Middlesex, England. Telephone: (0895) 36355 Telex: 934731
Belgium N.V. GIBCO Europe S.A. Drongensesteenweg 68, 9000 GENT, Belgium. Telephone: 091/277031 Telex: 12101
France GIBCO BRL SARL, 14 Rue des Oziers, BP 7050, 95051 CERGY-PONTOISE, France. Telephone: (1) 34-64-54-40 Telex: GIBBRL 609098F
Netherlands GIBCO Europe b.v., Hekven 3, 4824 AD BREDA, The Netherlands. Telephone: 76-41 99 44 Telex: 64016
Switzerland GIBCO BRL AG, Zwingerstrasse 12, CH 4018 BASEL, Switzerland. Telephone: 061 354066 Telex: 64520
Germany GIBCO BRL GmbH, Postfach 1212, Dieselstrasse 5, 7514 EGGENSTEIN, West Germany. Telephone: 721 705006 Telex: 7825957

Reader Service No.45

Pragmatism in public health

Alan Glynn

Inventing the NIH: Federal Biomedical Research Policy, 1887–1937.
By Victoria A. Harden.
The Johns Hopkins University Press:1986. Pp.274. $32.50, £25.50.

DAVID Smyth, the physiologist, once said "If primitive man had something on his mind it was probably because he had nothing in his stomach". The primitive hunger which led eventually to the founding of the National Institute of Health was simply for public health. The NIH's forerunner was the Laboratory of Hygiene set up in 1878 on Staten Island, New York, to detect infections among immigrants by using the new science of bacteriology. The Laboratory was part of the Marine Hospital Service, itself founded in 1798 to provide medical care for seamen.

Victoria Harden has drawn on numerous primary and secondary sources for her analysis of the forces which led to the establishment of the NIH in 1930. The author is not a scientist and her book is not a history of bench science or laboratory medicine. It is a history of the pressures, scientific and political, private and public, which shaped United States federal policy on biomedical research. In the end it was this policy which, after 50 years of struggle, allowed the change from a one-man, one-room laboratory to the first National Institute of Health. Development of federal policies over 50 years produced eleven Institutes spread over 317 acres at Bethesda in Maryland, employing 14,000 people and using or distributing 30 per cent of all American health research funds.

There was a false start early on. The value of "sanitary science" had become apparent during the Civil War and after the War attempts were made to enlarge and reform the Marine Hospital Service by, among others, John Shaw Billings, now best known as the initiator of the *Index Medicus*. Billings failed at first but an epidemic of yellow fever in the Mississippi valley led to the formation of a new National Health Board with him in charge. The Marine Hospital Service fought back and, starved of funds, the Board soon withered away. This set a recurring pattern. Reformers, and there were many, came up against not only the natural inertia of a conservative Congress or President but the opposition, open or covert, of the grandees of the Marine Hospital Service who were determined that whatever change had to come they were going to remain in charge.

Harden has, rightly, followed through particular lines of development rather than given a strictly chronological account. There is a list of the principal advances made by workers in the Laboratory of Hygiene, but for those who lack a detailed knowledge of American history a table of the main political events would have been more useful. The Laboratory flourished during the "progressive era" before the First World War. The War, and the work it demanded, stimulated scientists to think of major collaborative research projects embracing many disciplines, though there was still an instinctive distrust of the institutionalization of science.

The book outlines the uneasy and shifting alliances between scientists, health workers and reforming politicians, and helps us to understand an apparent paradox: how a country with a strong philosophy of self-help and freedom from government interference managed to create the largest and most successful laboratory ever devoted to biomedical research almost entirely on government money. Except for one or two brief references, for Harden Europe does not seem to exist. This is a pity, because it is fascinating to compare the events and philosophies she discusses with what was happening at roughly the same time on the other side of the Atlantic.

For example, in its early days the Laboratory of Hygiene was only authorized to deal with practical matters of infectious disease. Its early researches into pellagra or even water pollution were, strictly speaking, illegal uses of government money; indeed a separate appropriation for research was not made until 1907. Sir John Simon in England had faced similar problems 40 years before, when he imposed his own ideas on a vague remit from the Privy Council and eventually persuaded them to authorize, as he put it, "investigations, not necessarily connected with our practical business at the moment . . .". That was in 1865, and because governments everywhere "respond to pestilences" Simon was no doubt helped by the spread of cholera and cerebro-spinal fever in Europe and the arrival of yellow fever in Swansea. As a distinguished pathologist and surgeon, Simon received much help and backing from his professional colleagues whereas the early proponents of the NIH were sometimes in conflict with the American Medical Association.

Perhaps the closest parallel in Britain is the setting up of the Medical Research

The National Institutes of Health in the 1980s, showing only part of the Bethesda complex. On the right is Senator Joseph E. Ransdell, sponsor of the bill that transformed the Laboratory of Hygiene into the NIH. (*Pictures respectively courtesy of the NIH and from the book under review.*)

Council, the origins of which go back to the Local Government Board of 1871 and more immediately to the Medical Research Committee of 1913. Perhaps because of the tradition of Chadwick, Simon and Farr there seems to have been less worry in Britain than in the United States about the state getting mixed up in health and research, and the MRC was protected from direct ministerial interference at least until the Rothschild era. Although the NIH also achieved considerable scientific independence, Congress continues to try to influence its projects — as in the bill to set up an Institute of Nursing, vetoed last year by the President.

Despite the prosperity of the 1920s, progress in public health was difficult under the "new era" politics of Coolidge. Coolidge never mentioned the subject in a speech, and though treasury secretary Mellon was more favourably inclined he could rely on the director of the budget, Herbert Lord, to keep the purse strings tightly drawn. Lord combined a professional dislike of spending money, especially on research, with the Christian Scientist's belief in its futility.

By the late 1920s, the story takes on some of the inevitability of a Greek tragedy. Harden describes in detail the negotiations through which, after many years of struggle, Congress was persuaded to pass the Parker bill giving more freedom to the Public Health Service. Eight days later Coolidge vetoed it. Again there is a detailed but lively account of the haggling and manoeuvering in Congress. This time the Ransdell bill, which specifically set up the National Institute of Health, failed partly because of opposition from Parker who was smarting at his own defeat. At last Coolidge went and Lord retired. After more hectic activity, both bills passed both houses and were signed by President Hoover. It was 1930 and the Great Depression had already started.

Throughout this story there are many colourful and interesting characters. Yet I have the impression that the colour and the passion have been carefully filtered out in the interests of analytical clarity. This is unfortunate for the amateur historian or the general reader, but for those who want to influence public health policy the message comes through clearly. They must educate the public, the legislators and the administrators. And education here means both short- and long-term campaigns. It means using every effort to win over pressure groups, influencing the influential and converting the doubtful. Well-timed epidemics are always a help. In short this is an excellent, though rather dry, textbook for wets. □

Alan Glynn is Director of the Central Public Health Laboratory, Colindale, London NW9 5HT, UK, and a Visiting Professor at the London School of Hygiene and Tropical Medicine, University of London.

Signal studies in psychology

D. Aled Jeffreys

Human Brainwaves: The Psychological Significance of the Electroencephalogram. By Jacob Empson. *Macmillan, London/Stockton, New York:1986. Pp.126. Hbk £25, $50; pbk £7.95.*

In 1934, Adrian and Mathews confirmed and extended Berger's earlier but largely ignored discovery that small, continually varying, electrical signals could be recorded from the human scalp. These electroencephalograms (EEGs) have been extensively studied over the past 50 years but, even so, relatively little is known about their origins and significance. In contrast to their well-proven clinical applications, therefore, the "spontaneous" EEG signals have had only limited uses in research into normal brain function. They have nevertheless proved to be useful indicators of the level of consciousness and, in particular, of the different stages of sleep. There have also been encouraging results from experiments in which stimulus-induced changes in the EEG — the so-called event-related or evoked potentials (ERPs or EPs) — have been used to investigate cognitive and, to a lesser extent, sensory aspects of brain function.

Newcomers to such electrophysiological studies of the human brain are unfortunately confronted with a vast and specialized literature, often with a high technical content. Jacob Empson has tried to alleviate the problem by providing a simple, non-technical introduction to the use of the EEG and ERPs in cognitive psychology. The first two chapters include brief accounts of the history of the EEG and of modern encephalographic practice; descriptions of the typical characteristics of the EEG and their use in the assessment of states of consciousness; and discussion of various theories on the origins of the alpha rhythm. There are also individual chapters on event-related (and slow) potential studies of psychological processes, and on the EEG and localization of function in the brain. Two final chapters are devoted to the author's particular interests of normal and abnormal aspects of sleep.

As expected, the book does not provide detailed reviews of the literature but instead contains selective summaries of what the author judges to be the most important experimental findings. One can, in several cases, argue with some of his particular choices or omissions; and also perhaps complain that a short summary similar to that given in the section on event-related potentials was not provided in every chapter. But such shortcomings must be judged against the difficulty of the author's task. The quality of production is quite good except that the print on several pages of my review copy was slightly smudged.

The book is written in a simple and clear style, although occasional sentences have uncertain meaning and, in two cases (pp. 17 and 31), do not make sense at all, presumably because of missing lines or words. On the whole, however, it is easy to read and meets its main objective of providing a coherent and non-technical survey of the topic. *Human Brainwaves* will serve as a useful primer for psychology students and others who need a general introduction to non-clinical, psychological aspects of the EEG. For those who wish to progress further, each chapter has a selected bibliography of more comprehensive review papers. □

D. Aled Jeffreys is Senior Research Fellow in the Department of Communication and Neuroscience, Keele University, Staffordshire ST5 5BG, UK.

Material success

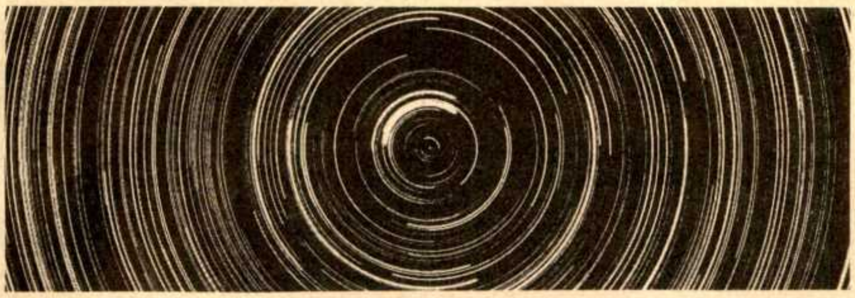
Charles Newey and colleagues

Encyclopedia of Materials Science and Engineering. Editor-in-chief Michael B. Bever. *Pergamon:1986. Eight volumes, pp. 6,015. £1,350, $1,950.*

The *Encyclopedia of Materials Science and Engineering* (EMSE) contains some five million words in seven volumes, plus a volume of indexes. How could a single reviewer possibly do justice to such a massive effort?

Depending on their stance inside or outside the "materials community", different users will look for different things from this encyclopaedia and one person could not hope to have a representative view. Further, review of EMSE requires assessment of a project planned by about 50 people and implemented by nearly 1,300 contributors writing over 1,500 entries. So, in keeping with the traditions of The Open University, we adopted a team approach. What follows is the essence of our corporate appraisal.

A good test of the usefulness of any encyclopaedia is to see how it performs against a list of topics you would expect or like to see covered. Our list included topics as diverse as Kyropoulos crystal growth, zeolites, charged dislocations, bubble memories, pyrometry, mercury reserves, globular proteins and value analysis. EMSE passed with flying colours, only failing on design for recycling (but recycling is well covered) and charged disloca-



The Earth moved — star tracks on an 8-hour exposure taken by a fixed camera pointed towards the North Pole; the boldest trail near the centre was made by Polaris. The picture is taken from the new paperback edition of Orbiting the Sun: Planets and Satellites of the Solar System *by Fred Whipple (reviewed in* Nature *294, 21; 1981). Publisher is Harvard University Press, price is $9.25, £6.75.*

tions (hardly surprising). In fact, the coverage is much more diverse than this with treatments of topics not conventionally associated with the books' title, for instance design, safety and health, materials and machines in antiquity, materials economics and management, art materials, newsprint and many materials of biological origin.

The Systematic Outline of EMSE (see Vols 1 and 8) groups the article titles under the 45 main subject headings used in the commissioning; they cover the production of materials, their nature and applications, property and structure assessment techniques, surface phenomena and processes, and more general topics such as mineral resources and management, and economics. These are further subdivided "as each subject dictates". The Introduction states that the space devoted to each of these sub-sections is in relation to its scientific, technical and industrial importance but we are given little insight into how this was assessed. The basis of selection must, in the end, be subjective and the updating policy will show whether the editors have met their goals.

The 1,570 articles are arranged alphabetically, which is a rather mixed blessing. This format should force users to look outside their specialities and, one hopes, help the editors to achieve their aim of cross-fertilization. The juxtaposition of "Titanium Alloys" with "Timbers from Various Geographical Locations" and "The Effect of Tax Upon Materials Policy" demonstrates this amply. On the other hand, the scheme puts a lot of stress on the cross-referencing system and it doesn't hold up too well, especially in terms of alternative titles (for example, "glassy" and "amorphous"). In addition, because the articles are not grouped under, say, specific materials, properties, processes or artefacts, the subject index — which is in Vol. 8 and very thorough — is an absolute must if you are looking for something specific.

Basically, EMSE consists of two types of entry: specific articles on particular topics in the science and engineering of materials, and articles of a more general kind. By and large the specific articles are authoritative, if succinct, treatments of their subjects but there is considerable variation in the level and approach. Perhaps this is inevitable, given the need to involve so many different authors each with their own ideas about the task, but there does seem to have been a certain unwillingness either to give clear briefing or to wield the editorial pen in the quest for consistency.

It is with the general articles that EMSE is clearly trying to be more than a conventional encyclopaedia. These are either reviews (for example "Electrical Materials — An Overview" and "Energy Considerations in Metals Production") or articles of an integrative nature ("Ceramics: Relation of Microstructure to Processing History", "Product Liability and Materials" and "Design and Materials Selection and Processing", for instance). The integrative articles are particularly informative and stimulating; they are good illustrations of the way to link aspects of materials that are often treated in isolation and, therefore, inadequately. Their inclusion is a clear manifestation of the editors' laudable desire to lower the barriers between disciplines. But beware, you will not find these gems in the subject index — for these you will need to use the Systematic Outline (Vol. 8).

In addition to unevenness in the level of the articles, the work can also be criticized for its slight bias towards the United States, and for a hint of partisanship. However, at this early stage in what should be an evolutionary process for EMSE, we will call this carping and trust that the balance is adjusted in the next edition.

The editors have tackled an enormous challenge and have met with considerable success. These volumes will be an important reference source for all scientists and engineers involved with the properties, processing and uses of materials, and especially for teachers of courses in the materials field. ☐

Charles Newey, Mark Endean, Peter Lewis, Nick Reid, Graham Weaver, George Weidmann and John Wood are members of the Materials Discipline, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

Living for physics

Michael Holwill

Introductory Biophysics. By M. Cerdonio and R.W. Noble. *World Scientific:1986. Pp.216. £20.25, $25.30.*

For those who browse among the shelves of libraries or bookshops, the title of this book and the attractive seascape on the dust-jacket will provide little indication of the contents. It is not an introductory text for students embarking on a course in biophysics; rather, it is aimed at scientists with a reasonable background in physics, particularly in thermodynamics, and it is intended to demonstrate that the basis of living systems can be understood and explained by the application of physical principles. This is not a new concept, but recent developments in both molecular cell biology and irreversible thermodynamics make it appropriate to present a modern account of the ideas concerned.

After a brief description of the characteristics by which living systems can be identified, and a chapter dealing with the structure and organization of those systems, from viroids through bacteria to multicellular organisms, the authors apply physical ideas to a number of biological areas before discussing, in the final chapter, the origins of life. The chapters are essentially independent, each containing some interesting features but having little direct relationship to the others. There is, in fact, much of importance and value in the book, but the rather dated diagrams (which vary considerably in style and quality) and the use of old-fashioned units (for example, "entropy units" instead of $J\ K^{-1}\ mol^{-1}$ and "calories" for joules) detract from the modern scientific approach.

It is convenient to have, collected together in a single volume, accounts of the essentials of various physico-chemical processes and techniques relevant to a comprehensive examination of the nature and development of living systems. By reducing the space occupied by some of the diagrams, however, a little more introductory material could have been added without the need for additional pages, and this would have increased the appeal of the work to the great number of scientists who have a general interest in this controversial topic. In the final chapter the authors provide some pertinent insights into how thermodynamics, chemical kinetics and fundamental particle physics may be of importance in specifying evolutionary processes, but they have not capitalized on the material earlier in the book by giving a clear summary of the current status of this analytical study. ☐

Michael Holwill is Reader in Biophysics in the Department of Physics, King's College London (KQC), Strand, London WC2R 2LS, UK.

That certain time in history

W.H. Brock

The Age of Science. By David Knight. *Basil Blackwell: 1986. Pp.251. £17.50, $24.95.*

THE reappraisal of the Victorians began in the 1940s. Much derided in the 1920s, today even politicians are reminting a specially-selected set of their "values" for the 1980s. Notwithstanding the material and cultural gaps between 1986 and 1886 or 1836, much of our present world can be clearly seen as derived from the Victorians — notably because of the vigorous export, through travel and emigration, of Victorian institutions and attitudes overseas, including parliamentary government, schools and universities, and the civil service, and religion and attitudes towards women and the family. Arguably, even the harmonies of contemporary popular music are those of Victorian church music. And, outside the relativistic and quantum boundaries of atomic physics and astrophysics, our vaunted twentieth-century science is essentially Victorian in origin.

In *The Age of Science,* David Knight, the most prolific of contemporary English historians of science (this is his eighth book), examines the scientific world-view in the nineteenth century. Although we may well believe that our present era is unprecedented in its domination by science, the reader will be persuaded by Knight that science equally dominated nineteenth-century thought and debates on religion, education, literature and the economic viability of the nation, though in a very different manner to our own. Here was a period before disciplines became so specialized that their practitioners became alienated from one another, when scientist, clergyman, writer and artist still spoke a common cultural language, and when science was ambitious, cocksure of progress and in its hegemonic role as the Church Scientific. As Knight points out, whereas historians may (wrongly) ignore the impact of science on earlier centuries, science cannot be left out of the narrative of Victorian history or of the historians' repertoire of explanations of what happened in history.

Knight writes deftly and entertainingly. Whether examining Anglo-French scientific rivalry, the challenge of German science, Victorian criteria for science and pseudoscience, the use and significance of scientific illustrations, symbols and jargon, the symbiosis of Victorian science and religion, or the slow erosion of scientific certitude, he wears his learning lightly. *The Age of Science,* a summary of the specialized reading Knight has done during the past 20 years, is a *tour de force.* It not only succeeds in making the history of science popular, but restores its 1950s role as a bridge between the cultures and disciplines which are the consequence of the Victorians' age of science. □

W.H. Brock is Director of the Victorian Studies Centre and Reader in the History of Science at the University of Leicester, Leicester LE1 7RH, UK.

Cold comforts

Beverly Halstead

How to Deep-Freeze a Mammoth. By Björn Kurtén. *Columbia University Press:1986. Pp.121. $16.96.*

A SCIENTIST'S general musings in a series of light essays on his subject are the dread of reviewers; few ever rise to the riveting heights of a Stephen Jay Gould. Björn Kurtén's slim volume, however, is well above the average. Kurtén specializes in the study of mammals of the Ice Age, and he has written fictional accounts of the life and times of the period.

Most of the essays deal with aspects of the Pleistocene, but also cover such topics as the third early bird *Archaeopteryx,* the early history of continental drift and the great Messinian salinity crisis when the Mediterranean dried out. This last event came to light when salt deposits 1.5–2 km in depth were discovered during the deep-sea drilling prograamme; after the drought came the flood, as the Atlantic poured in through the Straits of Gibraltar — the largest waterfall in Earth history.

It is, however, in writing about the Pleistocene that Kurtén has the surest touch — he does in fact explain how to deep-freeze a mammoth. It has happened but rarely. Normally the outer layers freeze but the deeper parts putrefy and all that remain are bones and bacteria. However, on a southern-facing slope in winter, it is possible for an entire animal to be freeze-dried, and Kurtén describes the discovery of a 30,000-year-old bison from Alaska which made a palatable stew.

There is, too, an essay on Piltdown Man; here, being a Finn Kurtén misses some of the subtlety of English humour. He refers to the "elephant club" from Piltdown that had been carved with a steel knife, but this unique artefact was seen by all, even the duped Sir Arthur Smith-Woodward in his book *The First Englishman,* as akin to a cricket bat. What else would the first Englishman require? The first European in the form of the now-notorious Petralona skull from Greece is featured, but in this case the discussion verges on the technical as Kurtén has been personally involved. In general, though, he has a light touch. Most of the essays are very readable, and the first and last of them are outstanding.

In the first, entitled "The Two Cultures", Kurtén contrasts the fundamental difference between physics and chemistry on the one hand, and biology on the other, expressing with great clarity the attitudes of many practitioners of the "woolly" sciences:

> But the world of physics is no more than a corner of the world of science. A man is greater than a star. A bacterium is greater than a star. A star cannot become very much else than a star — but bacteria have evolved into men .... Biology is the study of individuals, all of them unique, in inexpressible multiplicity and richness. And so the world of biology is something very different from the world of physics and calls for a different approach and different concepts.

The last essay is on cave art. It may be wondered if there could possibly be anything more to say that has not already been said many times before, but one surprising fact is that it appears that the Ice Age was a Golden Age: there was an abundance of food and the lifespan of human beings was not paralleled until virtually the present day. Furthermore, when cave art is examined from the standpoint of a zoologist an unusual perspective emerges: a male reindeer on heat or seals swimming upside down, there is a remarkable precision of observation. Kurtén suggests that many of the paintings may have been for children, in fact for their education. Interestingly, children's footprints are common on the floors of these caves. It would be important for them to recognize the animal's state and hence predict its behaviour.

When it comes to the portrayal of the human form, Kurtén points out that here too there is great accuracy of observation. He notes that many examples of cave art portray sexually inviting attitudes and are not dissimilar to photographs in pornographic magazines of the present day. The vulva is seen frontally, inverted and also from the back, open for penetration. Kurtén sees "unabashed sensual joy and tenderness" in all this. One has the distinct impression that palaeolithic man had a very great interest in and knowledge of sex, and he seems to have portrayed many species including his own in states of sexual arousal. Kurtén must surely be right; it must have been a Golden Age. □

Beverly Halstead is a Reader in the Departments of Geology and Zoology, University of Reading, Whiteknights, Reading RG6 2AH, UK.

- Recently published by the British Museum (Natural History) in the UK, and Facts on File in the United States, is *Mammal Evolution: An Illustrated Guide,* with text by R.J.G. Savage and illustrations by M.R. Long. The book will be reviewed in a future issue of *Nature.*

Cost-effectiveness of short-term tests for carcinogenicity

Lester B. Lave & Gilbert S. Omenn

Most chemicals to which we are exposed are not properly tested for carcinogenicity. The latest methods of* in vitro *testing provide a way of screening with sufficient accuracy to remedy this situation.

CONCERN about the dangers to health from exposure to toxic chemicals is increasing and all Western nations have developed elaborate schemes for regulating their production and use[1-6]. At present, decisions are made based primarily on long-term animal bioassays[1] which cost more than $1 million for each chemical and take 2-4 years. Here we present a framework for assessing the effective and efficient regulatory use of inexpensive, short-term *in vitro* tests and illustrate its application with data from an international collaborative testing programme[7].

The framework

Even though scientific data are rarely conclusive they provide the foundation for regulatory decisions[8] on whether a chemical should be regarded as carcinogenic or not at a particular exposure level. Carcinogenicity is particularly difficult to investigate because the causal mechanisms are far from fully understood, the period between exposure and onset of the disease may be decades, data on levels, routes, and time of exposure are generally poor, and susceptibility may vary due to predisposing genetic traits, pre-existing illnesses, personal habits such as cigarette smoking, and exposures to other chemicals that interact with the chemical of interest[9,10]. Ideally, current data would correctly classify all chemicals into true positives (TP) and true negatives (TN); unfortunately, it is inevitable that some misclassifications, false positives (FP) and false negatives (FN), will occur. Society should attempt to minimize the sum of the costs of testing and of misclassification.

In Fig. 1 we assume that either 2% or 10% of all chemicals are carcinogens. Measuring the cost of additional testing is straightforward, but what are the social costs of misclassification? It is not easy to generalize about the social costs of continued exposure to a carcinogen (FN). For example, lax standards for asbestos have resulted in hundreds of additional cancer cases each year. In contrast, lowering the benzene standard in the US from 10 to 1 p.p.m. in air was estimated to save less than one cancer per year. Respecting society's deep concern, we put a value of $10 million on a FN.

Estimating the social cost of a FP may be even harder, since it involves regulating a chemical needlessly. Capital and operating costs of exposure control equipment vary. If chemical production is discontinued, there is loss to the economy and substitutes have to be found. Cyclamates and the detergent nitrilotriacetic acid, both of which have been banned, are probably good examples of FPs. Substitutes may also turn out to have health risks, perhaps even greater ones. We put a cost of $1 million on a FP. For our purposes, the

a, Assume 2% (1,000) of 50,000 chemicals are carcinogenic

True biological activity

		Carcinogenic	Non-carcinogenic
In vitro test results	Potentially carcinogenic	TP 900	FP 4,900
	Not potentially carcinogenic	FN 100	TN 44,100

Social costs ($000 million)		
of false positives		
4,900 × $1 million	=	$4.9
of false negatives		
100 × $10 million	=	$1.0
Total	=	$5.9
Cost of testing		$0.4
Total social cost		**$6.3**
Social cost no testing, 1,000 FNs		
1,000 × $10 million		**$10.0**

b, Assume 10% (5,000) of 50,000 chemicals are carcinogenic

True biological activity

		Carcinogenic	Non-carcinogenic
In vitro test results	Potentially carcinogenic	TP 4,500	FP 4,500
	Not potentially carcinogenic	FN 500	TN 40,500

Social costs ($000 million)		
of false positives		
4,500 × $1 million	=	$4.5
of false negatives		
500 × $10 million		$5.0
Total	=	$9.5
Cost of testing		$0.4
Total social cost		**$9.9**
Social cost no testing, 5,000 FNs		
5,000 × $10 million	=	**$50.0**

Fig. 1 Social costs of testing chemicals for carcinogenicity. Illustrative analysis of consequences of FP (false positive) and FN (false negative) results, assuming that either *a*, 2% or *b*, 10% of chemicals are true carcinogens for humans, that the *in vitro* tests have a specificity and sensitivity of 90%, and that the cost of testing is $8,000 per chemical (test set 6).

critical assumption is that an FN costs ten times as much as an FP

As shown in Fig 1, there may be 1,000 or 5,000 unrecognized carcinogens among the 50,000 chemicals in widespread use The International Agency for Research on Cancer and the US National Toxicology Program list fewer than two dozen recognized human carcinogens[11,12] The current cost to society of FNs, therefore, may be $10,000 million to $50,000 million (at $10 million per carcinogen) These figures are disconcerting They are the hidden costs of an industrial society using chemicals with little knowledge about their potential carcinogenicity (or other adverse effects)

With a test system 90% accurate for both carcinogens and non-carcinogens and costing $8,000 per chemical, social cost would be $4,900 million for FP plus $1,000 million for FN (Fig 1*a*) Adding a testing cost of $400 million, social cost would be $6,300 million But if 2% of the 50,000 chemicals in use really are carcinogens, they have a social cost of $10,000 million The test system would lower this to $6,300 million If 10% are carcinogens, the test system would lower social cost from $50,000 million to $9,900 million

How high could misclassification be and still make the testing worthwhile, under these assumptions? If 2% of chemicals are carcinogens, the classification accuracy must not be lower than 84% for testing to be better than no testing If 10% are carcinogens, the classification accuracy need only be 52% Alternatively, if a slightly better test was available, one with 91% accuracy, it would be worth $11,800 or $19,000 more per chemical tested, assuming 2% or 10% of chemicals are carcinogens respectively

The above framework highlights the data required to make better social decisions about toxic chemicals the cost and accuracy of various carcinogenicity tests, the number of chemicals and proportion that are carcinogenic, and the social costs of false positives and false negatives Testing 50,000 chemicals in long-term animal bioassays is neither feasible nor cost-effective It is thus not surprising that increasing effort is being invested in the development and validation of many inexpensive, short-term tests *in vitro*[7,13–17] But misunderstandings remain about how accurate these tests would have to be to prove useful for regulatory decision making even though the 'gold standard', the long-term animal bioassay, itself has major limitations, both statistical and biological To be useful, a single short-term test need not be completely accurate, and the tests need not be useful for all classes of chemicals Short-term tests are inexpensive to replicate and provide information about dose-response relationships and mechanisms of action of chemicals, rather than simply counting the numbers of tumours in exposed animals

Table 1 Test chemicals carcinogen classification based on animal bioassay data[7]

Carcinogens	**Non-carcinogens**
4-dimethylaminoazobenzene (butter yellow)*	3,3′,5,5′-tetramethylbenzidine
chloroform*	4-dimethylaminoazobenzene-4-sulphonic acid Na salt
urethane*	pyrene
DL-ethionine*	α-butyrolactone
hexamethylphosphoramide (HMPA)*	anthracene
ethylenethiourea*	1,1,1-trichloroethane
diethylstilbestrol* (DES)	dimethylformamide
safrole*	diphenylnitrosamine‡
3-aminotriazole*	dinitrosopentamethylene tetramine
o-toluidine hydrochloride*	isopropyl-*N*-(3-chlorophenyl)carbamate
auramine (technical grade)*	L-methionine
4-nitroquinoline-*N*-oxide	sucrose
benzidine	L-ascorbic acid
benzo(*a*)pyrene	4-acetylaminofluorene†
β-propiolactone	3-methyl-4-nitroquinoline-*N*-oxide†
9,10-dimethylanthracene	1-naphthylamine†
2-naphthylamine	azoxybenzene†
dimethylcarbamoyl chloride	
N-nitrosomorpholine	
methylazoxymethanol acetate	
hydrazine sulphate	
cyclophosphamide	
epichlorhydrine	
4,4′-methylenebis-(2-chloroaniline)(MOCA)	
2-acetylamino-fluorene	

* 'Problem carcinogens'

† Classification changed from non-carcinogen to carcinogen as a result of the testing and excluded from the analysis

‡ Diphenylnitrosamine was re-classified when animal bioassay data became available, and is now included as a carcinogen

Test data

The short-term test framework can be illustrated with available test results The International Program for the Evaluation of Short-Term Tests for Carcinogenicity[7] was a collaborative study undertaken by the UK Medical Research Council/Health and Safety Executive, the US National Institute of Environmental Health Sciences and Environmental Protection Agency, and the Japanese National Cancer Center Research Institute We were not involved in the study and gratefully acknowledge the assistance of the program organizers

Test chemicals were selected to include examples from the major classes of widely used organic compounds Based on previous animal bioassay results, the 42 chemicals shown in Table 1 were classified as 25 carcinogens and 17 noncarcinogens, including 14 structurally similar carcinogen–noncarcinogen pairs and 11 'problem carcinogens' (reported to be difficult or impossible to detect in standard bacterial mutation assays) The chemicals certainly do not represent a random sample of all chemicals or all widely used chemicals After changes in classification during testing, a final analysis was based on 38 chemicals, 26 carcinogens and 12 non-carcinogens Compared with the US Environmental Protection Agency's Gene-tox data base, this base has proportionally more non-carcinogens[18], it excludes fibres, heavy metals, and some other agents known to cause problems in short- or long-term bioassays

Thirty assay systems (Table 2) were selected to represent a wide variety of biological endpoints and mechanisms of action but cytogenetic methods may have been under-represented Some investigators were funded by the study, while other investigators financed their own testing The investigators were not informed of the identity of the chemicals until after submission of their results Results were reported as positive, negative, or inconclusive[7] Unfortunately, quantitative results and data on strength of evidence were generally not reported Of the chemicals in Table 1, four (*o*-toluidine hydrochloride, HMPA, safrole, DES) are included in the *in vitro* assays and another four (benzo(a)pyrene, pyrene, 2-acetylamino-fluorene, 4-acetylamino-fluorene) are being subjected to intensive *in vivo* assays in collaborative follow-up studies[18]

Statistical method

We begin by deriving an estimate of the probability that a chemical is carcinogenic from the results of the short-term tests We define the dependent variable as taking the value of 0 when the chemical is believed to be a noncarcinogen and a value of 1 when it is believed to be a carcinogen Each independent variable is the report of a test on the chemical, it assumes a value of 0 for negative test results and 1 for positive tests Inconclusive results and chemicals not tested are excluded from the analysis

The relationship between dependent variable (presumed carcinogenicity) and

Table 2 Test classification accuracy based on collaborative test results

Test number	Type of test	Cost* ($)	$N^\dagger$	Accuracy
Bacterial mutation assays				
1	*S typhimurium*/plate	1,200	37	0 70
2	*S typhimurium*/plate	1,200	36	0 69
3	*S typhimurium*/plate	1,200	34	0 59
4	*S typhimurium*/plate	1,200	38	0 71
5	*S typhimurium*/fluctuation	NA	31	0 71
6	*S typhimurium*/plate	1,200	38	0 63
7	*S typhimurium*/plate	1,200	33	0 64
8	*S typhimurium*/plate	1,200	38	0 68
9	*S typhimurium*/plate	1,200	37	0 57
10	*S typhimurium*/plate	1,200	37	0 65
11	*S typhimurium*/plate	1,200	37	0 73
12	*S typhimurium*/plate	1,200	38	0 68
13	*S typhimurium*/plate	1,200	38	0 71
14	*S typhimurium*/+norharman	NA	28	0 57
15	*S typhimurium*/fluctuation	NA	33	0 55
16	*S typhimurium*/azaguanine-res	NA	38	0 66
17	*S typhimurium*/E coli WP2/fluc hepatocytes	NA	36	0 67
18	*S typhimurium*/*E coli*/ WP2/plate	NA	36	0 78
19	E coli WP2/plate	400	34	0 65
20	*E coli* 343	NA	38	0 72
Bacterial repair, phage induction, degranulation, nuclear enlargement assays				
21	*B subtilis*, M45 RecA-	800	38	0 79
22	*E coli*, RecA-/PolA	1,500	34	0 59
23	*E coli*, RecA-	1,500	37	0 62
24	*E coli*, RecA-/PolA	1,500	36	0 58
25	*E coli*, PolA	NA	38	0 59
26	λ-induction (*gal*⁺)	NA	29	0 65
27	λ-induction (lysis)	NA	36	0 56
28	Degranulation of RER	NA	33	0 39
29	Degranulation of RER+ ribonuclease	2,500	30	0 67
30	Nuclear enlargement, HeLa cells	NA	22	0 32
31	Nuclear enlargement, human fibroblasts	NA	22	0 64
Yeast assays				
32	*S cerevisiae* XV185-14C	NA	32	0 63
33	*S pombe* P1	NA	31	0 68
34	*S cerevisiae* PG-148, PG-154, PG-155, PG-166	NA	29	0 66
35	*S cerevisiae* D4	NA	35	0 43
36	*S cerevisiae* D6	NA	31	0 42
37	*S cerevisiae* D7	1,400	35	0 66
38	*S cerevisiae* JD1	NA	36	0 67
39	*S cerevisiae* rad	NA	37	0 59
***In vitro* mammalian test systems**				
40	UDS (unscheduled DNA synthesis) in fibroblasts	5,200	18	0 39
41	UDS in fibroblasts	5,200	21	0 62
42	UDS in HeLa cells	5,200	38	0 68
43	SCE (sister chromatid exchange)	3,000	19	0 53
44	SCE	7,500	18	0 67
45	SCE	3,000	33	0 52
46	CYT-VTRO	NA	21	0 57
47	CYT-VTRO	NA	18	0 72
48	mouse lymphoma L5178Y	4,900	19	0 53
49	CHO-HPRT	NA	9	0 44
50	CHO-HPRT	6,500	3	0 67
51	V79-HPRT	6,500	5	0 60
52	DIPHTH-R	NA	12	0 83
53	Transformation	NA	34	0 59
54	Transformation	NA	38	0 81
55	MLV-INFC	NA	24	0 52
***In vivo* assays**				
56	*Drosophila*-sex-linked recessive lethal	10,000	9	0 44
57	*Drosophila*-sex-linked recessive lethal	10,000	15	0 53
58	*Drosophila*-sex-linked recessive lethal	10,000	9	0 44
59	SCE-*in vivo*	3,000	17	0 59
60	Micronucleus	3,400	29	0 66
61	Micronucleus	3,400	17	0 47
62	Micronucleus	3,400	34	0 47
63	Sperm	11,400	15	0 47
64	Tradescantia	NA		

Collaborative test result from ref 7 NA, not available

* Estimated 1981 costs

† Number of chemicals tested in the international programme

independent variables is fitted to the testing results using logit analysis[19] The result is an estimated probability, P, that a chemical is carcinogenic If the analysis assigned $P=0$ to all noncarcinogens and $P=1$ to all carcinogens, the classification would be perfect With real data, a 'cut-point', P', is needed to classify all chemicals with $P<P'$ as noncarcinogens and all chemicals with $P>P'$ as carcinogens[20] As P' is increased, the number of FPs fall, but FNs appear The goal is to choose P' to minimize social costs $k\mathrm{FN}+\mathrm{FP}$, where k is the cost of an FN relative to that of an FP If the average social costs of FN and FP were identical ($k=1$), social cost would be minimized by having the smallest number of misclassifications However, if FN is valued higher than FP ($k=10$, as in Figs 1 and 2), then the cut-point is shifted to lower values (from B to A in Fig 2) In general P' is significantly greater than zero

In sum, the statistical procedure has two stages First, the probability of a chemical being a carcinogen is estimated by using observed short-term test results Second, a cut-point is sought that minimizes social cost in classifying chemicals

Reproducibility

Many of the short-term tests employed are relatively new and have not been standardized In the International Program, twelve investigators used versions of the *Salmonella*/microsome plate assay Eleven of the 38 chemicals were chosen because previous results in this type of assay system were unclear No two investigators reported identical results across the 38 chemicals, agreement between investigators ranging from 54% to 97%, even though the assay is standard Such discrepancies complicate interpretation and analysis[21,22]

Test batteries

Thirty short-term tests were examined in the International Program They comprise most (but certainly not all) of the short-term tests currently used or proposed for carcinogenicity testing Notably absent are DNA repair in cultured hepatocytes[23] and various bone marrow cytogenetic tests Every test was found to be of value in predicting carcinogenicity, with classification accuracies up to 81%, as shown in Table 2 Long-term bioassays may well offer no greater predictive accuracy, among 245 National Cancer Institute bioassays, 55 (22%) were judged inconclusive or equivocal[24]

While some tests show impressive

accuracy of classification, a battery of tests should be able to attain even greater accuracy, in large part because a variety of mechanisms can be screened Several proposals have been made concerning what combination would be best and these are discussed in Tables 2 and 3 Five previous proposals[25–28] as well as a set we drew from this data base are shown in Table 3 and the results of fitting the data to each test battery are shown in Table 4

Predicted probabilities of carcinogenicity for each of the 38 chemicals as estimated by test set 6 are shown in Fig 2 The estimated cut probability for $k=1$ would be between 0 42 and 0 81, for $k=10$, P' would be less than 0 05 The apparent superiority of test set 6 can be ascribed to our method of selecting tests These results indicate that test 21 (*Bacillus subtilis* Rec-assay) may deserve more attention than it has been given, despite current uncertainties about its mode of action

Value of testing

How much should society spend to acquire data before making a decision that a particular chemical should be treated as a potential carcinogen in humans? The more nearly complete and the more accurate the data, the lower the chance of an

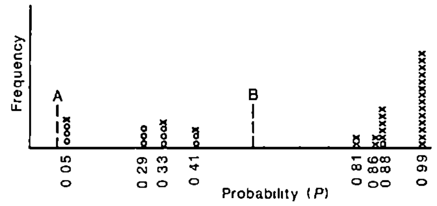


Fig 2 Predicted probability (P) of carcinogenicity This histogram for test set 6 shows the frequency of chemicals with P for the set of 38 chemicals in the International Collaborative Program Cut points corresponding to FN/FP values of 10(A) and 1(B) are indicated by dashed lines This cut point must be greater than zero to accommodate false positive test results × indicates a carcinogen while O indicates a noncarcinogen The three FN (from left to right) are chloroform, auramine, and DL-ethionine, the FP is α-butyrolactone

FN or FP The general principle is that data should be gathered until the time and dollar costs of acquiring more data approximate the additional savings to society from lowering the number of FP and FN Of course, the costs and savings are borne by different parties

The direct cost of additional data is the price of running more tests Each test has an associated accuracy and price Some tests are 'best buys' in the sense that no

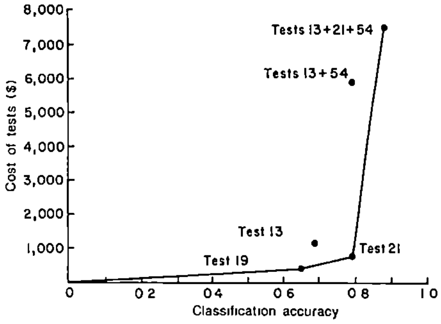


Fig. 3 Cost-effectiveness of *in vitro* testing The costs of tests are plotted against the estimate of test accuracy listed in Table 1 Tests 13, 21, and 54 correspond to set 6 in Table 2

test offers the same accuracy for a lower price Accuracy depends on the sample of chemicals tested Additional helpful tests give more information, especially when carcinogenic chemicals have heterogeneous mechanisms, but costs rise

The process of 'buying' additional classification accuracy can be represented schematically in Fig 3, where the vertical axis represents the cost of testing and the horizontal axis the classification accuracy of the tests in the International Program In this example, test 13 is both more expensive and less informative than test

Table 3 Test sets analysed

Set number	Test set suggested by	Assays suggested for inclusion	Versions tested
1	Bridges[25]	*Salmonella*/microsome Sex-linked recessive lethal in *Drosophila** Malignant transformation Mutation induction in cultured mammalian cells Yeast *Neurospora*†	*Salmonella*/microsome (test 13) Cell transformation, BHK-21 (test 54) Forward mutation, *S. pombe* (test 33)
2	W L Marcus, Office of Toxic Substances US EPA[26]	*Salmonella*/microsome *S pombe* Mammalian cell culture* Oncogenic transformation Unscheduled DNA synthesis in human cells	*Salmonella*/microsome (test 13) Forward mutation, *S pombe* (test 33) Cell transformation, BHK-21 (test 54) Unscheduled DNA synthesis, HeLa (test 42)
3	1980 National Toxicology Program Annual Plan[24]	*Salmonella*/microsome Mammalian cell culture Cell culture transformation	*Salmonella*/microsome (test 13) Unscheduled DNA synthesis, HeLa (test 42) Cell transformation, BHK-21 (test 54)
4	I F H Purchase, Imperial Chemical Industries Ltd (see ref 25)	*Salmonella* TA1535 & TA1538 *S cerevisiae** BHK-21 cell transformation Sebaceous-gland suppression† Rabin's test (degranulation)	*Salmonella*/microsome (test 13) Cell transformation, BHK-21 (test 54) Rabin's test (degranulation) (test 29)
5	Weisburger and Williams[28]	*Salmonella*/microsome Cell transformations, CHO mammalian mutagenesis Williams' hepatocyte† Sister chromatid exchange in CHO or liver cell systems*	*Salmonella*/microsome (test 13) Unscheduled DNA synthesis, HeLa (test 42) Cell transformation, BHK-21 (test 54)
6	Our analysis		*Salmonella*/microsome (test 13) Differential killing, *B subtilis*, (test 21) Cell transformation, BHK-21 (test 54)

The short-term tests in Table 2 are grouped into five categories by the mechanisms tested From the 3 to 13 available assays that represented each category, we constructed a new battery (set 6) by choosing the tests whose results in the International Program data set had explained the highest proportion of variance in the dependent variable (presumptive carcinogenicity) For all batteries except set 3, at least one test had to be omitted

* Assays included in International Program, but tested on too few chemicals

† Assay not included in International Program

Table 4 Logit regression equations

	Test set equation	No of chemicals	R^2	χ^2	Predictive accuracy C	Predictive accuracy NC
1	$A=-3\ 1+3\ 1T_{13}+2\ 8T_{33}+2\ 8T_{54}$	29	51 9	17 3	4/9	20/20
2	$A=-4\ 0+3\ 4T_{13}+2\ 8T_{33}+1\ 8T_{42}+2\ 8T_{54}$	29	55 3	19 3	5/9	20/20
3 & 5	$A=-2\ 8+2\ 3T_{13}+2\ 6T_{42}+3\ 2T_{54}$	38	48 1	22 6	5/12	26/26
4	$A=-1\ 8+43\ 9T_{13}+-0\ 7T_{29}+2\ 3T_{54}$	30	44 7	9 6	6/9	20/21
6	$A=-3\ 1+2\ 2T_{13}+2\ 4T_{21}+2\ 7T_{54}$	38	56 2	22 9	11/12	23/26

The general logit equation is $P=1$, where P is the estimated probability that a chemical is a carcinogen and A is a linear equation of estimated coefficients multiplied by test results[17] For example, in set 6, A is calculated as −3 1 plus 2 2 times the result of test 13 (either 0 or 1) plus 2 4 times the result of test 21 (0 or 1) plus 2 7 times the result of test 54 (0 or 1) These coefficients are estimated to give the best fit between the dependent variable, presumptive carcinogenicity, and the P calculated from this logit equation The coefficient of test 54 (2 7) indicates that this test is more important in predicting carcinogenicity than the other two tests (coefficients 2 2 and 2 4) The omission of some chemicals from the data set and the relatively small size of the sample appear to have led to the nonsense result that test 29 in set 4 should be interpreted as the opposite of the test results R^2 should be interpreted as the proportion of the variation in the dependent variable accounted for by the regression, the χ^2 statistic shows that the test sets were all statistically significant Also shown are the proportions of carcinogens (C) and noncarcinogens (NC) (actually tested) that were classified correctly All the sets perform well, giving classification accuracies up to 89%, although the test batteries selected to be useful in predicting carcinogens are notably less successful in predicting noncarcinogens Test set 6 correctly predicts 10 of the 11 problem carcinogens known to be difficult for some of these tests to handle '

21 The combination of tests 13 and 54 gives more information at greater cost than 21 alone, the combination of tests 13, 21, and 54 provides more information at only a slight increase in cost The solid line shows the tests and combinations of tests that provide classification accuracy at lowest cost For this sample of chemicals, if only about $1,000 per chemical were available, test 21 would provide 79% classification accuracy If about $7,400 per chemical were available, tests 13, 21, and 54 could provide 89% classification accuracy

Because the classification accuracy of the various test batteries is about 90% and the total cost of testing 50,000 chemicals, with these batteries estimated at $8,000 per chemical, would be $400 million, our Fig 1 is not purely hypothetical It seems clear that society is making a serious mistake in failing to test most chemicals

Testing all chemicals immediately is neither feasible nor intelligent A better approach is a sequential testing policy with orderly selection of priority chemicals for testing (based on production, exposures, and structural similarity to known carcinogens) and use of information on metabolism and mechanisms of action to select tests and interpret results[29-32] Evaluation of initial test results will lead to modifications of the test regime and classification criteria Tier-testing systems can address inconclusive or conflicting results The above discussion is simplistic in dichotomizing chemicals as carcinogenic or harmless, rather than estimating the quantitative toxicity or potency, as testing becomes more sophisticated, it may be possible to estimate potential risks quantitatively

Conclusion

We have presented a decision analysis framework for the testing and classification of toxic chemicals, using an extraordinary set of published data The available set of tests and chemicals is small and difficult to interpret, but it is the best available and illustrates the framework The set of tests that together predicted the carcinogenicity of the 38 chemicals tested most successfully may not be appropriate for other chemicals, much less the entire population of 50,000 chemicals in use While the specific results strictly hold only for this limited data set, several general conclusions can be stated

First, short-term tests can be accurate predictors of the carcinogenicity of chemicals Sets of tests provided classification accuracy of 89% even when 11 of the 38 chemicals were known to be difficult to detect as mutagenic in the *Salmonella*/microsome assay

Second, decision analysis modelling of classification problems in terms of the social costs of false positives and false negatives can be instructive If the social costs are equal, classification accuracy should be maximized When the social costs of a false negative are greater than those of a false positive, the cut-point should be lowered to classify more chemicals with intermediate probability estimates as carcinogens

Finally, it may be cheaper for society to test all chemicals in common use with batteries of short-term tests, despite the resulting false positives and false negatives, rather than to continue tolerating thousands of unregulated carcinogens because there are no animal bioassay test results Several sets of short-term tests appear to be sufficiently accurate and inexpensive to warrant expanded use in testing However, a more intelligent procedure than routinely testing all 50,000 chemicals would be to test first the highest priority chemicals, based on production and exposures, and to evaluate the test batteries and classification criteria as testing progresses The US Environmental Protection Agency has authority under the Toxic Substances Control Act to demand a regime of tests for new chemicals that meet one of two Section 4 criteria (1) an unreasonable risk of injury to health or the environment based on its physical or biochemical characteristics, or (2) expected production in substantial quantities, giving rise to significant human or environmental exposures.

We recommend that the development and validation of short-term tests for application in screening new and existing chemicals should be accelerated

---

Lester B Lave is James H Higgins Professor of Economics in the Graduate School of Industrial Administration, Carnegie-Mellon University, Gilbert S Omenn is Dean of the School of Public Health and Professor of Medicine and Environmental Health, University of Washington

We thank Kathleen D Heffernan and Glen Dranoff This work was begun when the authors were at the Brookings Institution Lave's work was supported by the NSF, Omenn's work by the Sloan and Charles A Dana Foundations and the National Cancer Institute

1 Lave, L B *The Strategy of Social Regulation* (Brookings, Washington, 1981)
2 Lave, L B *Quantitative Risk Assessment for Regulation* (Brookings, Washington, 1982)
3 Calkins, D R, Dixon, R L, Gerber, C B, Zarin, D & Omenn, G S *J nat Cancer Inst* **64**, 169-176 (1980)
4 Ruckelshaus, W D *Issues in Science and Technology* **1**, 19-38 (1985)
5 Weinstein, M C *Public Policy* **27**, 333-343 (1979)
6 *Regulations Relating to Environmental Chemicals* (OECD, Paris, 1977)
7 deSerres, F J & Ashby, J (eds) *Evaluation of Short Term Tests for Carcinogens Report of the International Collaborative Program* (Elsevier, New York, 1981)
8 National Research Council *Risk Analysis in the Federal Government Managing the Process* (National Academy Press, Washington, 1983)
9 Friedman, R D *Sensitive Populations and Environmental Standards A Legal Analysis* (Conservation Foundation, Washington, 1979)
10 Omenn, G S *J occup Med* **24**, 369-374, 1982
11 IARC Monographs on the *Evaluation of the Carcinogenic Risk of Chemicals to Humans Chemicals, Industrial Processes and Industries Associated with Cancer in Humans* (International Agency for Research on Cancer, Lyon, 1982)
12 National Toxicology Program *Third Annual Report on Carcinogens* (US Department of Health and Human Services, Washington, September, 1983)
13 Ames, B N *Science* **204**, 587-593 (1979)
14 Office of Science and Technology Policy, *Fed Register* **49** (100), 21594-21661 (1984)
15 Purchase, I F U *Mutat Res* **99**, 53-71 (1982)

16 National Toxicology Program *Review of Current DHHS, DOE, and EPA Research Related to Toxicology* (FY 1983) (US Department of HHS, Washington, 1983)
17 Shelby, M D & Stasiewicz, S , *Envir Mutagen.* **6**, 871-876 (1984)
18 Waters, M D , Garrett, N E , Covone de Serres, C M , Howard, B E & Stack, H F *Genetic Toxicology of Some Known or Suspected Human Carcinogens,* in *Chemical Mutagens Principles and Methods for their Detection* Vol 8 (ed de Serres, F J) 261-341 (Plenum, New York, 1983)
19 Goldstein, M & Dillon, W R *Discrete Discriminant Analysis* (Wiley, New York, 1978)
20 Lave, L B , Omenn, G S , Heffernan, K D & Dranoff, G *J Am. Col Toxicol* **2**, 125-130 (1983)
21 Chankong, V , Haimes, Y Y , Rosenkranz, H S & Pet-Edwards, J *Mutat Res* **153**, 135-166, 1985
22 Pet-Edwards, J , Chankong, V , Rosenkranz, H S & Haimes, Y Y *Mutat Res* **153**, 187-200, 1985
23 Williams, G D , Laspia, M F & Dunkel V C *Mutat Res* **97**, 359, 1982
24 *National Toxicology Program Tech Bull* **1**(3), 2 (1980)
25 Bridges, B S in *Screening Tests in Chemical Carcinogenesis* (eds Montesano, R , Bartsch, H , Tomatis, L) 549-568 (International Agency for Research on Cancer, Lyon, 1976)
26 Marcus, W L in *Carcinogenesis, A Comprehensive Survey* Vol 3 *Polynuclear Aromatic Hydrocarbons* (eds Jones, P W & Fruendenthal, R. I) 469-472 (Raven, New York, 1978)
27 Office of Technology Assessment *Assessment of Technologies for Determining Cancer Risks from the Environment* 135-136 (US Government Printing Office, Washington, 1981)
28 Weisburger, J H & Williams, G M *Science* **214**, 401-407 (1981)
29 Hoel, D G , Kaplan, N L & Anderson, M W *Science* **219**, 1032-1037 (1983)
30 Crump, K. S , Hoel, D G , Langley, C H & Peto, R *Cancer Res* **36**, 2973-2979, 1976
31 Omenn, G S in *Molecular and Cellular Approaches to Understanding Mechanisms of Toxicity* (ed Tashjian, A H Jr) 224-245 (Harvard School of Public Health, Boston, 1984)
32 National Research Council *Toxicity Testing* (National Academy Press, Washington, 1984)

ARTICLES

Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene

Kirk R. Thomas & Mario R. Capecchi

Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

Injection of homologous DNA sequences into nuclei of cultured mammalian cells induces mutations in the cognate chromosomal gene It appears that these mutations result from incorrect repair of a heteroduplex formed between the introduced and the chromosomal sequence This phenomenon is termed 'heteroduplex induced mutagenesis' The high frequency of these events suggests that this method may prove useful for introducing mutations into specific mammalian genes

GENE-TARGETING (homologous recombination between DNA sequences residing in the chromosome and newly introduced DNA sequences) allows the specific alteration of genes in the mammalian genome[1-4] Recently, Smithies *et al*[1] introduced a DNA sequence into the chromosomal β-globin locus by homologous recombination We have corrected mutant neomycin resistance genes (*neo*) in the host genome via homologous recombination with an injected plasmid DNA carrying a different mutation in the *neo* gene[2] In the process of analysing these gene-targeting events, we uncovered an unexpected class of 'corrected' chromosomal *neo*^{r} genes, which retained the original mutation but had acquired a second mutation that restored gene function The acquisition of the compensating mutation apparently occurred as a result of incorrect repair of a heteroduplex formed between the *neo* gene in the chromosome and the newly introduced *neo* gene These reactions occur at a surprisingly high frequency In addition, we show that the *neo*^{r} genes containing the compensating mutations function in mammalian cells and in *Escherischia coli* as a result of a novel translation reinitiation mechanism

Recombinant plasmids and recipient cell lines

We first established cell lines containing an amber mutant *neo* gene integrated into the genome of the mouse fibroblast line LMtk$^-$ We then corrected this mutation by injecting plasmid DNA carrying a non-overlapping deletion mutation in the *neo* gene Cell lines containing a corrected *neo*^{r} gene were identified by selecting for cells resistant to the drug G418

Figure 1*a* shows the recombinant plasmids, pRH4-14/TK and pRH140ΔNae/TK, used for these experiments These plasmids were derived from the parental plasmid pRH140[5], which contains sequences from the bacterial plasmid pBR322 and the *neo*^{r} gene coded by the bacterial Tn*5* transposon The pBR322 sequences supply an ampicillin resistance gene (*amp*^{r}) and an origin of DNA replication which functions in bacteria The *neo*^{r} gene was engineered to be functional both in bacteria and in mammalian cells[5] In bacteria, the *neo*^{r} gene confers kanamycin resistance, in mammalian cells the *neo*^{r} gene confers resistance to the drug G418 (G418^r) The herpes simplex virus thymidine kinase gene, HSV-*tk*, was introduced into the above plasmid at the unique *Bam*HI site to generate the plasmid pRH140/TK

pRH4-14/TK contains an amber codon near the 5′ end of the *neo* gene[5,6] This premature polypeptide chain termination signal renders the gene product defective in both bacteria and mammalian cells, the mutation also creates a new *Dde*I site which is used as a test for the presence of the amber mutation pRH140ΔNae/TK contains a deletion of 284 bp at the 3′ end of the *neo* gene which removes 52 amino acids from the carboxy-terminal end of the NEO protein, rendering it non-functional The plasmid pRH4-14-ΔNae/TK contains both the 4-14 amber mutation and the ΔNae deletion

We used two recipient mouse cell lines derived from LMtk$^-$ fibroblasts Both cell lines habour the plasmid pRH4-14/TK, but differ in the copy number and chromosomal location of the integrated plasmid[2] The cell line LM1 contains a single copy of the plasmid integrated into the host genome by its *Hin*dIII ends at a locus designated *J*-1 Cell line LM4 contains four copies of pRH-14/TK integrated at four independent chromosomal sites designated *J*-1-*J*-4

Generation of G418^r cell lines

Plasmid DNA, (~5 molecules per cell) was injected into nuclei of LM1 and LM4 cells, which were then selected with G418 Injection of linear pRH140ΔNae/TK molecules into LM1 or LM4 resulted in G418^r cell lines at a frequency of ~1 per 1,000 cells injected (Table 1) This frequency is five orders of magnitude greater than the spontaneous reversion frequency[2] (to G418^r of either cell line) No G418^r cell lines were obtained following injection of LM1 or LM4 with pBR322/TK, pRH4-

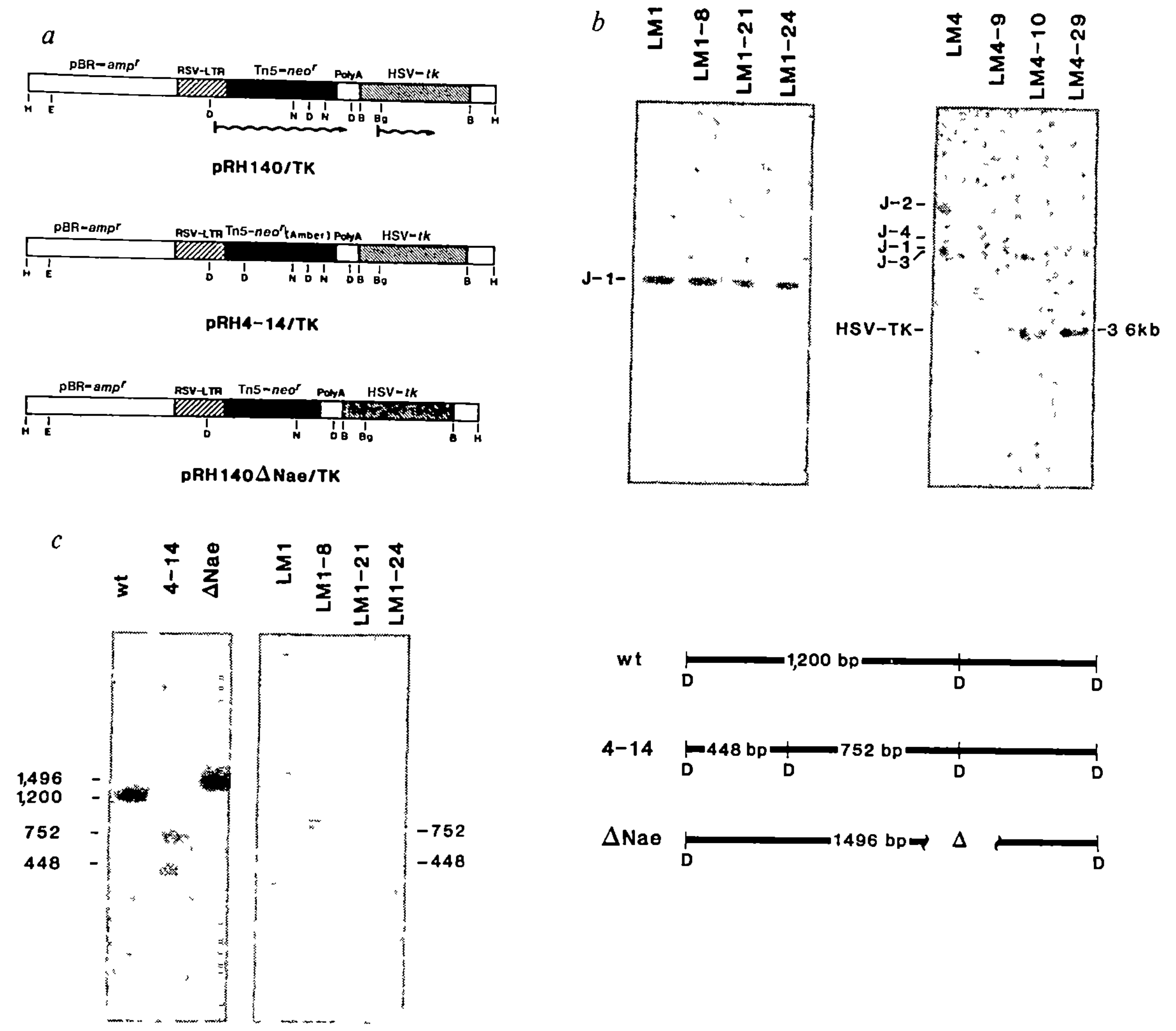


Fig 1 *a*, Maps of the plasmids pRH140/TK, pRH4-14/TK and pRH140ΔNae/TK *b*, *Bgl*II and *Bam*HI Southern transfer patterns of LM1, LM4 and their G418^r derivatives *c*, *Dde*I Southern transfer pattern of LM1, LM1-8, LM1-21 and LM1-24 *a*, The plasmids contain sequences derived from pBR322, the *neo*^{r} gene coded for by the bacterial Tn*5* transposon and the herpes simplex virus thymidine kinase gene (HSV-*tk*) The *neo*^{r} gene is expressed from a bifunctional promoter, RSV-LTR, which allows expression of the *neo*^{r} gene in *E coli* and in mammalian cells[5] pRH4-14/TK contains an amber mutation that also creates a new *Dde*I site pRH140ΔNae/TK contains a 284 base pair deletion at the 3′ end of the *neo* gene Each vector is represented in linear form from the unique *Hin*dIII site The restriction sites are designated H, *Hin*dIII, E, *Eco*RI, Bg, *Bgl*II, N, *Nae*I, B, *Bam*HI and D, *Dde*I Poly A, polyadenylation sequence from HSV-*tk*. For a more extensive restriction map of these plasmids see refs 2, 6 *b*, DNA was purified from each cell line and digested with either *Bgl*II (LM1, LM1-8, LM1-21, LM1-24) or *Bam*HI (LM4, LM4-9, LM4-10, LM4-29) Aliquots of DNA (5 μg) were electrophoresed through 0 75% agarose, transferred to nitrocellulose paper and probed with either ^{32}P labelled, nick-translated pRH140 (LM1) or pRH140/TK (LM4) kb, kilobases *c*, DNA (10 μg) from each cell line was digested with *Dde*I, electrophoresed through 0 75% agarose, transferred to nitrocellulose and probed with a ^{32}P-labelled, nick-translated, 1200-bp *Dde*I fragment from the *neo*^{r} gene of pRH140 The sizes (in base pairs) of restriction fragments generated by *Dde*I digestion of the three *neo*^{r} alleles (wild type, the 4-14 amber mutation and the ΔNae deletion mutation) are given

14/TK or pRH-14-ΔNae/TK (see Table 1)

Two classes of G418^r cell lines were obtained from LM1 and LM4 In the first class a wild type *neo*^{r} gene was generated by homologous recombination between the newly introduced pRH140ΔNae/TK plasmid sequence and the pRH4-14/TK chromosomal sequence This class has been described previously[2] The second class became resistant to G418^r by virtue of 'heteroduplex induced (het-induced) mutagenesis'

Fig 1*b* shows the Southern transfer patterns of LM1 and LM4 and several of their G418 derivatives Digestion of genomic DNA with either *Bam*HI or *Bgl*II will isolate the integrated *neo* genes onto single fragments linked to chromosomal sequences In the case of LM1, *Bgl*II digestion reveals a single hybridizing band representing the single *neo* locus, *J*-1 The *Bgl*II Southern transfer patterns of the G418^r derivative cell lines LM1-8, LM1-21 and LM1-24, are indistinguishable from the parental pattern Digestion of LM4 DNA with *Bam*HI reveals 4 bands which hybridize to pRH140 sequences *Bam*HI digestion of DNA from the G418^r derivatives LM4-9, LM4-10 and LM4-29 show identical patterns Thus conversion to G418^r was not the result of acquisition of new *neo* sequences or rearrangements of old ones

As shown in Fig 1*c*, digestion of genomic DNA with the restriction enzyme *Dde*I generates a series of fragments that allow the detection of the wild-type *neo*^{r} gene and each of the *neo* mutant alleles, the 4-14 amber mutation and the ΔNae deletion mutation The Southern transfer patterns of three G418^r cell lines derived from LM1 following digestion of genomic DNA with *Dde*I are also shown Only the *Dde*I fragments characteristic of the 4-14 amber mutation are seen Similar results were obtained when DNA from LM4 and its derivative G418^r cell lines, LM4-9, LM4-10 and LM4-29 were digested with *Dde*I These results were quite unexpected because all the derivative cell lines were G418^r

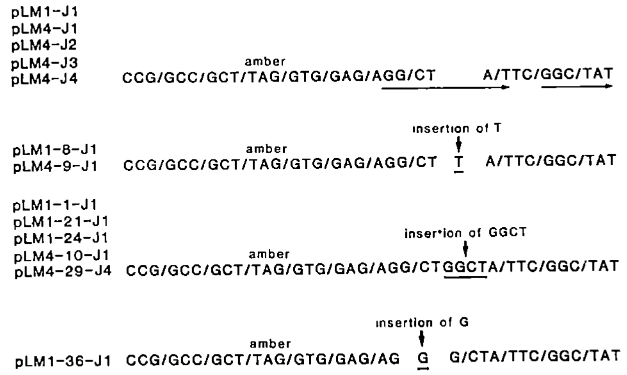

pLM1-J1
pLM4-J1
pLM4-J2
pLM4-J3
pLM4-J4
amber
CCG/GCC/GCT/TAG/GTG/GAG/AGG/CT A/TTC/GGC/TAT
insertion of T
pLM1-8-J1
pLM4-9-J1
amber
CCG/GCC/GCT/TAG/GTG/GAG/AGG/CT T A/TTC/GGC/TAT
pLM1-1-J1
pLM1-21-J1
pLM1-24-J1
pLM4-10-J1
pLM4-29-J4
insertion of GGCT
amber
CCG/GCC/GCT/TAG/GTG/GAG/AGG/CTGGCTA/TTC/GGC/TAT
insertion of G
amber
pLM1-36-J1
CCG/GCC/GCT/TAG/GTG/GAG/AG G G/CTA/TTC/GGC/TAT

Fig 2 A summary of pertinent DNA sequences from plasmids rescued from LM1 and LM4 and their G418^r derivatives Plasmids were rescued from each cell line as described[2] 5′ fragments of the *neo*^{r} genes from these plasmids were subcloned into the M13 vector, mp8, and sequenced by the chain termination method[8] The sequences of the protein coding strands representing codons 12–22 are shown The amber mutations as well as the 6 base pair repeats are underlined

Plasmids were rescued from the genome of LM1, LM1-8, LM1-21 and LM1-24 For these experiments genomic DNA was digested with *Bam*HI or *Bgl*II and ligated under conditions that favoured intramolecular ligation This DNA was used to obtain ampicillin resistant bacteria by transfection[6] From each cell line we rescued a class of plasmids containing the same 5′ junction with chromosomal DNA Plasmids derived from LM1, as expected, do not provide resistance to kanamycin, but the plasmids rescued from the G418^r cell lines do When we examined the *Dde*I polymorphism present in these rescued plasmids, all had the 4–14 *Dde*I polymorphism including the plasmids rescued from LM1-8, LM1-21 and LM1-24 Thus all the plasmids retained the amber mutation, yet those rescued from the G418^r cell lines conferred kanamycin resistance on bacteria

Plasmids were also rescued from LM4, LM4-9, LM4-10 and LM4-29 after digestion of genomic DNA with either *Bam*HI or *Bgl*II One plasmid from each G418r line was found to confer kanamycin resistance on *E coli* When these plasmids were reintroduced into LMtk$^-$ cells G418^r colonies were obtained at a frequency similar to that resulting from injection of the wild-type *neo*^{r} gene Thus these plasmids contain a *neo*^{r} gene that functions in both bacteria and mammalian cells The rescued plasmids all had the 4-14 *Dde*I polymorphism, including those that conferred drug resistance (*J*-1 from LM4-9 and LM4-10 and *J*-4 from LM4-29)

Table 1 Transformation frequencies to G418^r by injecting LM1 and LM4 cells with plasmid DNAs

DNA injected	Frequency of G418^r cell lines generated by hetero-duplex induced mutagenesis	Number of cells injected
pBR/TK	0	2×10^4
pRH140ΔNae/TK	$1\ 2\times10^{-3}$	10^4
pRH4-14/TK	0	3×10^4
pRH4-14-ΔNae/TK	0	2×10^4

LM1 or LM4 cells were grown on glass cover slips (10 mm × 10 mm) in 35 mm Petri dishes Twenty-five cells per dish received nuclear injections of ~5 plasmid molecules per cell[7] Plasmid molecules were linearized with *Hin*dIII (pBR/TK) or *Bcl*I (pRH140ΔNae/TK, pRH-14/TK, pRH4-14-ΔNae/TK) After the injection, the cells were incubated for 24 h in nonselective medium at 37° in a 5% CO_2 incubator and then switched to minimum essential medium supplemented with 400 μg ml^{-1} G418 The dishes were scored for colonies after 3 weeks pBR/TK contains the 3 6 kb *Bam*HI fragment of the herpes *tk* gene inserted at the *Bam*HI site of pBR322 pRH4-14/TK and pRH140ΔNae/TK are described in Fig 1 legend pRH4-14-ΔNae/TK was created by the deletion of the 284 bp *Nae* fragment from the *neo*^{r} gene in pRH4-14/TK

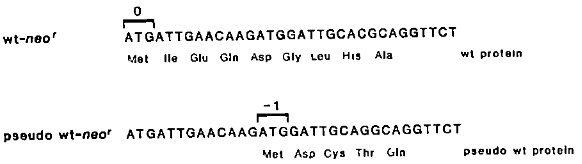

0
wt-neo^r
ATGATTGAACAAGATGGATTGCACGCAGGTTCT
Met Ile Glu Gln Asp Gly Leu His Ala
wt protein
-1
pseudo wt-neo^r
ATGATTGAACAAGATGGATTGCAGGCAGGTTCT
Met Asp Cys Thr Gln
pseudo wt protein

Fig 3 The amino-terminal sequences of wild-type and pseudo-wild-type *neo*^{r} proteins are shown with the DNA sequence of the 5′ end of the *neo*^{r} gene To facilitate protein purification, hybrid proteins were made by in-frame fusions of the 5′ ends of *neo*^{r} genes with the *lacZ* gene Wild-type fusions were made with the *neo*^{r} gene from the plasmid *p*RH140 (ref 5), pseudo-wild-type fusions were made with the *neo*^{r} gene from pLM1-1-J1 Both fusions were constructed by inserting a blunt-ended *Bcl*I–*Pvu*II restriction fragment, containing nucleotides −31 to +233 of *neo*^{r}, into a blunt-ended *Apa*I site in the 8th codon of the *lacZ* gene in p93 94 (R Weiss, unpublished) *E coli* cells containing the hybrid genes were grown to stationary phase, lysed in a French Press, and the extract clarified by centrifugation at 280,000*g* for one h The β-galactosidase activity was purified either by adhesion to antibodies (mouse anti- β-galactosidase, Promega Biotech), followed by acrylamide gel electrophoresis, or by affinity chromatography[9] The sequence of the first seven amino acids of each protein was determined on an ABI microprotein sequenator

The compensating mutations are insertions

We sequenced ~300 bp at the 5′ ends of the *neo*^{r} genes isolated from the G418^r cell lines These were compared with the corresponding sequences from the plasmids isolated from the parental cell lines LM1 and LM4 As illustrated in Fig 2, the only changes found in the *neo*^{r} genes were the insertion of a thymidine residue 11 bp downstream from the amber mutation in pLM1-8-J1 and pLM4-9-J1, the insertion of four bases, GGCT, in pLM1-1-J1, pLM1-21-J1, pLM1-24-J1, pLM4-10-J1 and pLM4-29-J4, and the insertion of a guanosine residue 8 or 9 bp downstream from the amber mutation in pLM1-36-J1

Reinitiation of translation

To determine how *neo*^{r} genes containing both an amber mutation and a frameshift mutation confer *neo*^{r} activity in both bacteria and mammalian cells, we analysed the amino-terminal sequence of the *neo*^{r} gene products We will refer to the gene containing both the amber mutation and an insertion mutation as the pseudo-wild-type gene To facilitate analysis, the 5′ 200 nucleotides of the *neo* genes were fused to the *lacZ* gene Hybrid β-galactosidase containing the amino termini from the wild-type and pseudo-wild-type NEO protein containing the amber mutation and the GGCT insertion were purified from transformed *E coli* and subjected to amino-terminal sequencing The fusion protein from wild-type *neo*^{r}-*lacZ* gene has the sequence Met-Ile-Glu-Gln-Asp-Gly as predicted (Fig 3) However, the protein from the pseudo-wild-type *lacZ* gene fusion began with the sequence Met-Asp-Cys-Thr-Gln This protein was initiated from an AUG codon 14 bp downstream from the normal AUG codon in the −1 translation reading frame (see Fig 3) It now becomes clear how the insertion mutations reverse the amber mutation After initiation of protein synthesis in the −1 frame, the ribosome passes through the amber mutation, which is in the 0 reading frame and therefore not read, and regains proper phase when it reaches the +1 frameshift mutation (insertion of a T, G or GGCT) Although this alters the amino-terminal amino acid sequence of the NEO protein, this portion of the protein has been shown to be dispensable[10]

Figure 4*a* shows two models that could explain how translation of the pseudo-wild-type gene begins at the −1 AUG codon

In model I we assume that the ribosome can enter at two sites, AUG codons in the 0 and −1 reading frame Model II uses a single entry point followed by translation reinitiation In this model the ribosome enters at the AUG codon in the 0 reading frame and translates the messenger RNA until it reaches the amber codon There it terminates and releases the polypeptide fragment The ribosome then scans back until it reaches the −1 AUG codon, where it reinitiates protein synthesis

We performed two sets of experiments to distinguish between these models First, the amber mutation was removed, by *in vitro* site-directed mutagenesis, from the pseudo-wild-type gene Second, a four bp insertion mutation was created *in vitro* in the wild-type *neo*^{r} gene just downstream from the −1 AUG codon Plasmids containing these altered pseudo-wild-type and wild-type genes were introduced into *E coli* and cultured mammalian cells to evaluate their ability to confer kanamycin and G418 resistance (see Fig 4*b*) The two models make opposite predictions of the resultant phenotypes If the ribosome can enter at either AUG then both the above plasmids should be functional, the ribosomes entering at the −1 AUG codon would regain proper phase on reaching either of the +1 frameshift mutations However, neither set of plasmids would be functional if only the initial entry site is used because (1) after entering the pseudo-wild-type mRNA lacking the amber mutation, the ribosome would be unable to terminate and therefore unable to reinitiate protein synthesis, (2) after entering a wild type mRNA containing a four base-pair insertion, the ribosome would encounter a +1 frameshift and generate a mutant gene product As shown in Fig 4*b*, neither of these *neo* genes is functional in either bacteria or mammalian cells These results strongly support the single entry-termination-reinitiation model

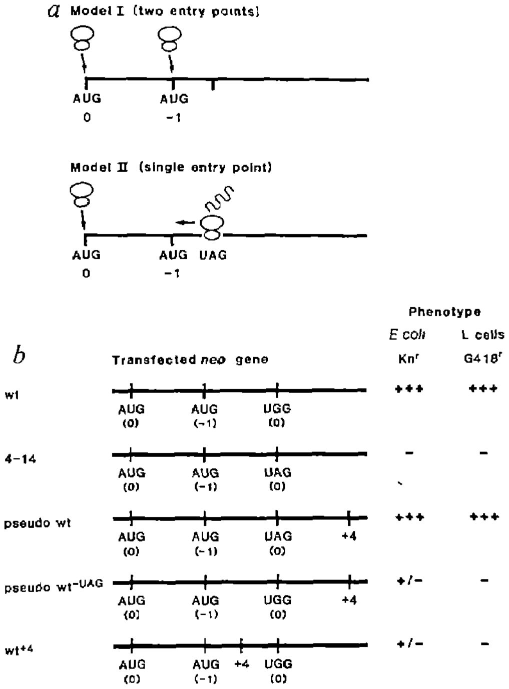


Fig 4 *a* Two models explaining the synthesis of pseudo-wild-type protein *b*, Phenotypic analysis of transfected *neo*^{r} genes *a*, In model I, ribosomes entering *neo*^{r} messenger RNA bind at two binding sites, one on the 0 reading frame at nucleotide 1, the other on the −1 reading frame at nucleotide 14 Proteins made from the latter site would be shifted onto the appropriate reading frame by downstream +1 frameshifts Model II proposes a single ribosome entry site in the 0 reading frame at nucleotide 1 To use the −1 AUG codon at position 14, the ribosome must first translate in the 0 reading frame up to the UAG (amber) codon At this point, the ribosome will release the newly synthesized peptide and then scan the message until it locates the alternative AUG codon in the −1 reading frame *b*, Plasmids containing various *neo*^{r} genes were assayed for the ability to confer drug resistance to *E coli* or to mouse L cells The wild type *neo*^{r} gene (wt) from pRH140 contains a tryptophan codon, UGG, at codon 15 4–14 is the amber mutation from the plasmid pRH4–14 and contains UAG at codon 15 Pseudo-wt is a sequence containing both the 4-14 amber mutation and a downstream, four base-pair insertion of GGCT, following nucleotide 56 Pseudo-wt$^{(-UAG)}$ was derived from the pseudo-wt gene, by removal of the UAG codon by site-directed mutagenesis[11] The mutant wt^{+4} was derived from the wild-type gene following the insertion of four base pairs at position 35 GGCC was inserted at this position by filling in the ends of an *Xma*III restriction site The figure depicts the 5′ end of the *neo*^{r} mRNA from each plasmid AUG (0), the translational initiation codon in the 0 reading frame, AUG (−1), the proposed initiation codon in the −1 reading frame at nucleotide 14, the sites of insertional mutagenesis of GGCC following base 35 or GGCT following base 56 are indicated (+4), UGG and UAG are the alternative states of the fifteenth codon Plasmids containing these mutations were introduced into *E coli* strain MH1 by CaCl$_2$-mediated transformation or into mouse L cells by nuclear microinjection Sensitivity of *E coli* to kanamycin or of L cells to G418, was determined as previously described[6] +/−, Resistance to low levels of kanamycin We believe that this low level of resistance was again due to reinitiation of the −1 AUG as these +1 frameshift mutations generate a new nonsense mutation in the +1 reading frame much further downstream from the − AUG (a UGA codon 53 bp downstream from the −1 AUG codon)

Discussion

The unexpected finding that correction of the pRH4–14/TK sequences in the host genome of LM1 and LM4 often results from insertion of a few base pairs downstream from the amber mutation raises a number of questions How did the insertions of these base pairs occur? How did injection of pRH140ΔNae/TK into either LM1 or LM4 mediate these events? The DNA sequence surrounding the position of the base-pair insertions contains the six bp direct repeat GGCTAT (see Fig 2) The fact that each of the insertions, T, G or GGCT is a tandem repeat of part of this sequence suggests a duplication-generating mechanism

A number of observations argue that the insertion or duplication events at this site were induced by the initiation of recombination between the pRH4–14/TK sequence residing in the chromosome and the incoming pRH140ΔNae/TK sequence First, the frequency of generating this class of G418^r cell lines was comparable to the frequency of generating cell lines by legitimate gene replacement or gene conversion[2] Second, this frequency is five orders of magnitude greater than the spontaneous reversion frequency of LM1 or LM4 to G418^r Third, we previously isolated a cell line in which both a homologous recombination event and a het-induced mutagenic event occurred[2] Because the homologous recombination events and the mutagenic events each occur at a frequency of ~1 per 1000 cells injected, the predicted frequency of both events occurring independently in the same cell line is 1 per 10^6 injected cells As we obtained such a cell line after a few thousand injections, it is tempting to postulate that the two events occurred as a result of a concerted reaction Fourth, we did not obtain G418^r cells from LM1 or LM4 following injection of a plasmid DNA, pBR322/TK, that does not contain sequences similar to the *neo* gene, so injection of DNA does not *per se* induce rampant mutagenesis Similarly, we did not obtain G418^r cells from LM1 or LM4 following injection of pRH4–14/TK or pRH4–14ΔNae/TK These experiments demonstrate the specificity of the reaction Because injection of the pRH4–14/TK vector does not induce the mutations, an important factor for triggering the mutagenic response may be the single base-pair mismatches at the amber mutation and/or the large mismatch at the ΔNae deletion The observation that injection of the double mutant, pRH4-14-ΔNae/TK, also does not induce mutations argues that at least the single base-pair mismatch is required The only mechanism that we can envision by which single base-pair mismatches between a chromosomal sequence and an exogeneous sequence can influence the mutation process requires formation of a heteroduplex

A number of models can be drawn in which the strands in the heteroduplex transiently misalign at the direct repeats leading to partial duplication of this sequence as a consequence of DNA repair A model mechanism for the insertions found in the LM1 and LM4 G418^r cell lines is given in Fig 5 A misaligned heteroduplex between the *neo* gene in the chromosome and the injected *neo* gene is formed, a nick is then introduced near one of the loops, generating a primer for strand extension The condition that the insertion mutation produces a functional gene restricts both the position of the nicks and the number of nucleotides that can be incorporated at the nick We have shown that insertions that generate a frameshift mutation in the +1 translation reading frame will compensate for the upstream amber mutation This condition limits the number of bases incorporated at the nick to one base, 4 bases, 7 bases etc The position of the nicks is also restricted Nicks at some positions followed by insertion of one or four bp introduce new nonsense mutations Thus, only a limited set of possible mutations exist which generate a functional *neo*^{r} gene After primer extension, the DNA strands are ligated and the insertion fixed into the genome by a round of DNA replication In prokaryotic recombination and/or DNA repair, reactions homologous to the ligation depicted in the model may not occur, however, such reactions do occur in mammalian cells[12]

We stress that all the insertion mutations that we analysed resulted from independent events They involve insertions of different bases, in different cell lines, at different chromosomal loci and were generated at different times It is interesting that all eight insertions occurred in the first repeat The proximity and/or the nature of the mismatch at the amber mutation may account for this polarity

The translation reinitiation mechanism described in this paper could be a general mechanism for translating polycistronic messenger RNAs in eukaryotic cells After the termination of the first protein, the ribosome could scan, forward or backwards, for the AUG codon used to initiate translation of the second protein Recent reports[13,14] support such a model for translation of a polycistronic messenger Further, the *src* mRNA of the Rous sarcoma virus and some genes of the cauliflower mosaic virus may be translated by such a mechanism[15,16] A similar mechanism can also explain the observation[17,18] made with *in vitro* engineered templates, that translation-initiation at an internal AUG codon is stimulated by the presence of a termination codon in frame with the upstream AUG codon

Conclusion

As a consequence of correcting a gene residing in the host genome by 'gene-targeting', we uncovered an unexpected class of corrected genes These genes still retained the original mutation but acquired a compensating mutation The frequency of this event was surprising and was shown to depend not only on homology but also on mismatched base pairs between the newly introduced DNA sequence and the corresponding sequence residing in the genome All the insertion mutations occurred in a small direct repeat proximal to the base-pair mismatch, suggesting that the direct repeat is also an important component for generating high frequency 'het-induced mutagenesis' The actual frequency of mutagenesis may be much higher than the observed frequency of 1 per 1000 cells receiving an injection because we imposed the condition that only mutations that converted the amber mutant into a functional gene will be identified Most such mutations would not be expected to restore gene activity

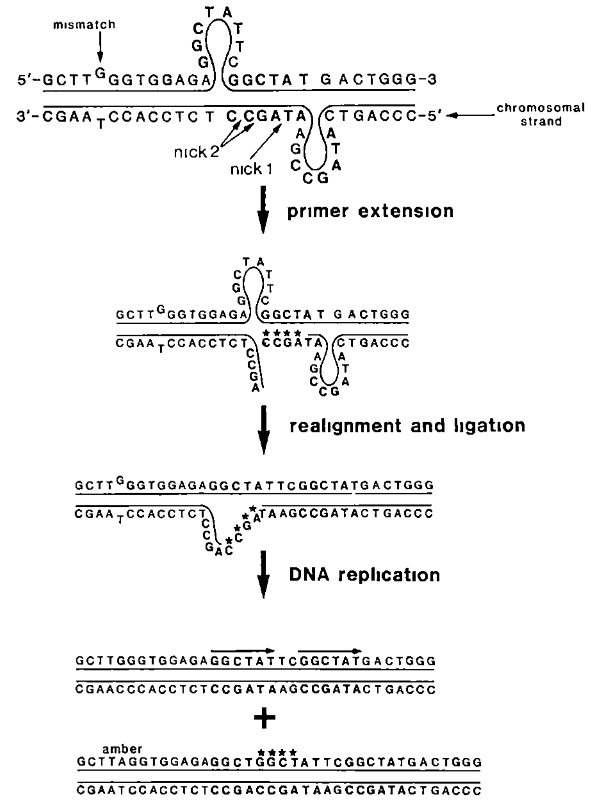

mismatch
5'-GCTT^{G}GGTGGAGA GGCTA T G ACTGGG-3
3'-CGAA$_T$CCACCTC T CCGATA CTGACCC-5'
chromosomal strand
nick 2
nick 1
primer extension
realignment and ligation
DNA replication
GCTTGGGTGGAGAGGCTATTCGGCTATGACTGGG
CGAACCCACCTCTCCGATAAGCCGATACTGACCC
+
amber
GCTTAGGTGGAGAGGCTGGCTATTCGGCTATGACTGGG
CGAATCCACCTCTCCGACCGATAAGCCGATACTGACCC

Fig. 5 A model for the mechanism of inserting GGCT into the chromosomal pRH-14/TK sequence The first step features the formation of a misaligned heteroduplex at the six base-pair direct repeat, between the *neo*^{r} gene in the chromosome and the injected *neo*^{r} gene In the second step a nick is introduced near one of the displaced loops to generate a primer for strand extension Following extension, the strands are ligated and the insertion fixed into the genome by a round of DNA replication The positions of the nicks were chosen to account for the insertions isolated nick 1 would generate either the +T or the +GGCT insertions, nick 2 would generate the +G insertion The site of the nick and the length of strand extension is limited by the requirement to produce a functional gene Bold letters, the bases involved in the six base-pair repeats, *, the bases added during strand extension

In future, we will examine 'het-induced mutagenesis' that directs the loss of gene function rather than the correction of gene function These events should occur at a higher frequency and reveal a wider spectrum of changes at the DNA sequence level The only apparent requirements for a 'hot spot' for het-induced mutagenesis are small direct repeats proximal to base-pair mismatches between the newly introduced plasmid sequence and the homologous sequence in the genome Comparable small direct repeats are encountered in the DNA coding sequence of most genes, making them susceptible to this type of mutagenesis Permutations of this methodology should provide the means for efficiently introducing mutations into specific mammalian cellular genes

We thank Laurie Fraser for assistance with tissue culture and Regina Zeikus, Robert Weiss and Diane Dunn for assistance with protein sequencing

Received 18 July, accepted 2 October 1986

1 Smithies, O, Gregg, R G, Boggs, S S, Koralewski, M A & Kucherlapati, R S *Nature* **317**, 230-234 (1985)
2 Thomas, K R, Folger, K R & Capecchi, M R *Cell* **44**, 419-428 (1986)
3 Lin, F L, Sperle, K and Sternberg, N *Proc natn Acad Sci USA* **82**, 1391-1395 (1985)
4 Smith, A J H & Berg, P *Cold Spring Harb Symp Quant Biol* **49**, 171-181 (1984)
5 Hudziak, R *et al Cell* **31**, 137-146 (1982)
6 Folger, K R, Thomas, K R & Capecchi, M R *Molec cell Biol* **5**, 59-69 (1985)
7 Capecchi, M R *Cell* **22**, 479-488 (1980)
8 Sanger F, Nicklen, S & Coulson, R *Proc natn Acad Sci USA* **74**, 5463-5467 (1977)
9 Ullman, A *Gene* **29**, 27-31 (1984)
10 Beck, E, Ludwig, G, Auerswald, E A, Riess, B & Schaller, H *Gene* **19**, 327-336 (1982)
11 Hutchinson, C A *et al J biol Chem* **253**, 6551-6560 (1978)
12 Roth, D B, Porter, T N & Wilson, J H *Molec cell Biol* **5**, 2599-2607 (1985)
13 Peabody, D S & Berg, P *Molec cell Biol* **6**, 2695-2703 (1986)
14 Peabody, D S, Subramani, S & Berg, P *Molec cell Biol* **6**, 2704-2711 (1986)
15 Hughes, S *et al Molec cell Biol* **4**, 1738-1746 (1984)
16 Dixon, L K & Hohn T *EMBO J* **3**, 2731-2736 (1984)
17 Johansen, H, Schumperli, D & Rosenberg, M *Proc natn Acad Sci USA* **81**, 7698 (1984)
18 Liu, C, Simonsen, C C & Levinson, A D *Nature* **309**, 82-85 (1984)

Outbursts of Cygnus X-3 observed at 1.3 and 3.3 mm wavelengths

J. W. M. Baars*, W. J. Altenhoff*, H. Hein† & H. Steppe†

* Max-Planck-Institut fur Radioastronomie, D-5300 Bonn 1, FRG
† Instituto de Radioastronomia Millimetrica, Aven divina Pastora 7, 18012, Granada, Spain

The X-ray source Cyg X-3 is associated with a radio source which shows irregular, fast and extremely strong outbursts of radiation. The first such event was seen in September 1972 by many observers over a wide frequency range[1]; several more events have occurred since then. The main characteristics of the outbursts have been described by the behaviour of an expanding synchrotron source, which undergoes adiabatic and synchrotron losses[1,8]. However, the detailed physics of the outbursts remains to be understood, in spite of a growing set of observational data. Here we report observations of Cyg X-3 made simultaneously at wavelengths of 1.3 and 3.3 mm, showing outbursts with an intensity of several Janskys and a duration of only a few hours.

Cyg X-3 is a binary object with a 4.8-h periodicity in both X-ray and infrared emission[10]. This periodicity may also be visible in low-level radio flares[2], but was not observed in several outbursts at a level above 50 mJy (ref. 3). Following a suggestion by K. Johnston that a flare was imminent, we included Cyg X-3 as a test object for continuum observations, during the commissioning of the new 30-m-diameter Millimetre Radio Telescope[4] of the French–German Institute for Radioastronomy in the Millimetre Range (IRAM).

Apart from an observation during the September 1972 outburst at 3.3 mm (ref. 5), no data are available on the development of a burst at millimetre wavelengths. Our observations made simultaneously at two wavelengths (1.3 and 3.3 mm) enable us to study the time dependence of the spectral index. We observed Cyg X-3 from 1 to 11 October 1985. On 1/2 October we barely detected Cyg X-3, whereas on 5/6 October it was easily seen. Subsequently, we followed the object on two consecutive days, from 17.00 UT on 6 October until 02.00 UT on 7 October, and from 16.00 UT on 7 October until 02.45 UT on 8 October. On 11 October a single observation at 3.3 mm was obtained.

The IRAM 30-m telescope is located at 2,850 m altitude on Pico Veleta in the Sierra Nevada near Granada, Spain. The telescope has an altazimuth mounting with a Nasmyth focus in a spacious instrument cabin. During our observations, the beams of the two receivers were coincident to within 2 arc s on the sky. The receivers were equipped with cooled Schottky-mixers and had a double-sideband system temperature (excluding noise from the antenna and the atmosphere) of 360 K and 200 K at the wavelengths of 1.3 and 3.3 mm (frequencies of 226 and 90 GHz), respectively, and a bandwidth of 500 MHz. Calibration of the system sensitivity is achieved with the aid of loads at ambient and liquid-nitrogen temperatures, introduced in the plane between the quaternary Nasmyth mirror and the receiver input.

At millimetre wavelengths the sensitivity is usually limited by fluctuations in the lower atmosphere. These can be effectively suppressed by quickly moving the antenna beam on and off the source and measuring the difference in the signal received from the two directions. In the lower region of the atmosphere, the two beams have an almost perfect overlap, so that the fluctuations are nearly cancelled in the difference signal[6].

For these 'beam-switched' observations we installed a focal-plane chopper. A vane in the form of a butterfly with a diameter of 1.5 m, rotated parallel to and in front of the quaternary mirror, moved the beam through an angle of 82 arc s on the sky at a rate of 6 Hz. Excellent suppression of atmospheric fluctuations was achieved. The half-power width of the antenna beam is 26 arc s at 3.3 mm wavelength and 12.3 arc s at 1.3 mm, and during our observing period, the aperture efficiencies at these wavelengths were 0.53 and 0.24, respectively.

Our method of observation consisted of measuring the ratio of the intensities of Cyg X-3 and the planetary nebula NGC7027. The angular distance between these two sources is only a few degrees, so that any dependence on the atmospheric extinction and antenna efficiency is eliminated from the ratio. NGC7027 is strong enough to be used for measuring pointing corrections, and, moreover, does not show any time variations in its flux density. We used Mars as a flux density calibrator. On 9 November 1985 we determined the ratio NGC7027/Mars, when Mars was near aphelion and had an apparent diameter of 4.04 arc s. We assumed a black-body brightness temperature of 198 K with an absolute error of probably <5% (ref. 7), and derived flux densities of 93.8 and 14.9 Jy, leading to observed peak flux densities for NGC7027 of 4.27 and 4.80 Jy at 1.3 and 3.3 mm, respectively.

Following a pointing check on NGC7027, its intensity was determined by a set of five on-off pairs, in which the on- and off-axis beams were alternately placed on the source for an integration time of 10 s each. Then a similar set of ten on-off pairs was obtained for Cyg X-3. Such a cycle lasted for 30 min, after which it was repeated.

Figure 1 shows the derived flux densities of Cyg X-3 as a function of time. The noise in a set of ten on-off measurements (200 s of integration time) is ~0.1–0.2 and 0.04 Jy at 1.3 and

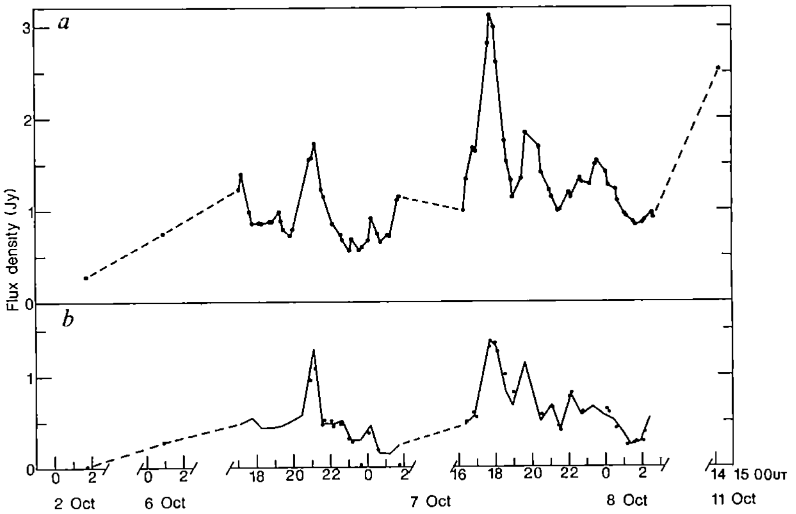


Fig. 1 Flux density of Cyg X-3 during 2–11 October 1985 at wavelengths of 3.3 (*a*) and 1.3 (*b*) mm, observed with the IRAM 30-m telescope. In *b* the line connects the averages of two adjacent measurements, which are shown individually as dots.

3 3 mm, respectively By combining pairs of data points the 1 3 mm results were smoothed considerably

During our observations three strong outbursts occurred almost simultaneously at both wavelengths The first and last are nearly equally strong at both frequencies (spectral index $\alpha = -0\ 2$, where $S \propto \nu^{\alpha}$ for flux density S and frequency ν), whereas the middle outburst is twice as strong at 3 3 mm ($\alpha = -0\ 7$) The average plateau-level flux density also has $\alpha = -0\ 7$ on both days

During the observation period the flux density of Cyg X-3 was monitored at 11 cm wavelength with the Green Bank interferometer (K. J Johnston, personal communication) Our first observation on 1/2 October coincides with the beginning of the precursor burst at 11 cm, and the main flare at 11 cm rose from 2 0 Jy at 00 00 UT on 6 October to 9 0 Jy at 00 00 UT on 8 October Our millimetre-wavelength observations show only a slow increase of the plateau level during this interval

The time development of the 11 cm flare is similar to earlier outbursts observed at wavelengths between 2 and 21 cm The first widely observed flare of September 1972 has been well described by a model of an expanding cloud of relativistic particles, mixed with an ionized thermal gas[8], in which the flare reaches its maximum earlier at shorter wavelengths

This model is clearly not appropriate for the character of the flare at millimetre wavelengths The duration of the millimetre-wavelength bursts is only ~1 h between the half-power points, with a suggestion of a broadening at 3 3 mm This could explain why these single, short outbursts are not seen at the longer centimetre wavelengths

The millimetre-wavelength bursts are probably related to a repeated injection of relativistic particles Such a scheme is necessary to adequately describe the early phase of the centimetre-wavelength flares[9] At injection the electrons will have the highest energy, and in losing their energy during the expansion, they will radiate first at the shortest wavelengths The millimetre-wavelength flares of short duration may be manifestations of separate bursts of injected particles, in which case differences in the optical depth of the medium could cause the observed differences in spectral index To understand how these short bursts build up to the observed centimetre-wavelength flare will require the analysis of detailed, simultaneous observations at millimetre and short centimetre wavelengths

We have looked in these data for the 4 8-h periodicity present in the X-ray and infrared emission of Cyg X-3 A spectral analysis indicates a possible periodicity of ~5 h in the 3 3-mm data on 6/7 October, which can be seen also in Fig 1 There is a suggestion of a 2 5-h periodicity in the 1 3-mm data on 7/8 October Because our observations on 7 October started almost exactly 3×4 8 h after we stopped observing on 6/7 October, we also combined the data and searched the total 20 h of data for a 4 8-h periodicity without finding a significant result We conclude that our data do not produce evidence for a 4 8-h periodicity in the millimetre-wavelength outbursts on 6–8 October 1985, similar to the result in ref 3 for centimetre-wavelength outbursts above 50 mJy Our noise level is too high to detect a possible periodicity in the low-level radiation[2]

We thank K J Johnston for a timely warning of a possible outburst and for data before publication The observation on 11 October was made by A Baudry

Received 6 May, accepted 31 July 1986

1 Gregory, P C *et al Nature* **239**, 440–443 (1972)
2 Molnar, L A, Reid, M J & Grindlay, J E *Nature* **310**, 662–665 (1984)
3 Johnston, K J *et al. Astrophys J* (submitted)
4 Baars, J W M., Hooghoudt, B G, Mezger, P G & de Jonge, M J *Astr Astrophys* (submitted)
5 Pomphrey, R. B, Epstein, E E *Nature phys Sci.* **239**, 125–126 (1972)
6 Baars, J W M *Nature* **212**, 494–495 (1966)
7 Ulich, B L *Astr J* **86**, 1619–1626 (1981)
8 Marscher, A P & Brown, R. L *Astrophys J* **200** 719–726 (1975)
9 Peterson, F W *Nature* **242**, 173–177 (1973)
10 Parsignault, D R *et al. Nature Phys Sci* **239**, 123–125 (1972)

Io plasma torus and the source of jovian kilometric radiation (bKOM)

Dyfrig Jones

Space Plasma Physics Group, British Antarctic Survey, NERC,
Madingley Road, Cambridge CB3 0ET, UK

One of the many discoveries by Voyagers 1 and 2 during their flyby of Jupiter was the detection of radio emission at kilometric wavelengths[1,2] in the frequency range 20–500 kHz. Broadband jovian kilometric radiation bKOM was observed to be beamed away from the zenomagnetic equatorial plane, the beaming angle increasing with increasing frequency[1–3]. It was suggested that bKOM is created by the linear conversion of electrostatic upper hybrid waves to electromagnetic waves in the density gradients at the Io torus[4]. The beaming is inherent to the conversion mechanism and depends on the characteristic frequencies of the source plasma Here I present results, based on this theory, of remote sensing of the Io torus using bKOM observed by Voyager 2 when inbound towards Jupiter. The density profiles obtained are in remarkable agreement with measurements of the torus densities obtained *in situ* by Voyager 1. This seems to provide very strong evidence that bKOM is produced at the Io torus by the linear mode-conversion mechanism and indicates that the technique used for remote sensing could be a powerful method of determining the torus characteristics when no *in situ* measurements are available.

bKOM first appeared in the planetary radio astronomy (PRA) data when Voyagers were >2 AU from the planet[1,2] Figure 1 shows a PRA wave spectrogram[5] in which bKOM is the tapered emission recorded when Jupiter's north magnetic pole (located at central meridian longitude, CML = 202°) is tilted towards the spacecraft The leading edges of the tapered emissions occur when Voyager 2 is at CML <202° and the trailing edges when Voyager's CML is >202° Voyager 2's jovicentric latitude when inbound (at 09 30 h jovian local time, LT), was 7 2° and, because Jupiter's magnetic dipole axis is tilted at 9 6° to its rotation axis, the spacecraft's excursion in magnetic latitude during one jovian rotation was −2 4° to 16 8° Spectrograms such as that in Fig 1 show that bKOM was generally observed when Voyager was at magnetic latitudes >10 °N Voyager 1, whose magnetic latitude varied between −6 5° and 12 7° when inbound (at 10 30 LT), and which therefore went to more southerly latitudes than Voyager 2, occasionally observed a similar beaming in the southern hemisphere[6] This would indicate approximate symmetry about the magnetic equator with a wedge-shaped region of no illumination near the equatorial plane, the width of the region depending on frequency Contrary to earlier reports[5] the radiation is right-hand polarized in the northern hemisphere and left-hand polarized in the south[6] The tapered form of the emisson evident in Fig 1 clearly shows that the higher frequencies are beamed to higher latitudes

The linear conversion theory[4] proposed for the production of bKOM invoked electrostatic upper hybrid (ESUH) waves as the source, with their propagation in the Io plasma torus ultimately resulting in the emission of the electromagnetic O mode radiation Intense ESUH waves were observed within 23 R_J of the planet which were confined to within 1 9° of the magnetic equatorial plane[7] In the theory, it is proposed that these waves propagate in the torus density gradient which is assumed to be normal to the magnetic field The ESUH waves first become electromagnetic Z mode waves[8] which, in turn, after further propagation, convert into electromagnetic O mode radiation through a radio window[9] which exists where the wave frequency equals the plasma frequency $f_p = 9 \times 10^{-3} N_e^{1/2}$ kHz, N_e m^{-3} being the electron density On passing through the radio window into the low-density region outside the torus the O mode radiation emanates as two beams at angle $\beta = \arctan\ (f_p/f_c)^{1/2}$ to

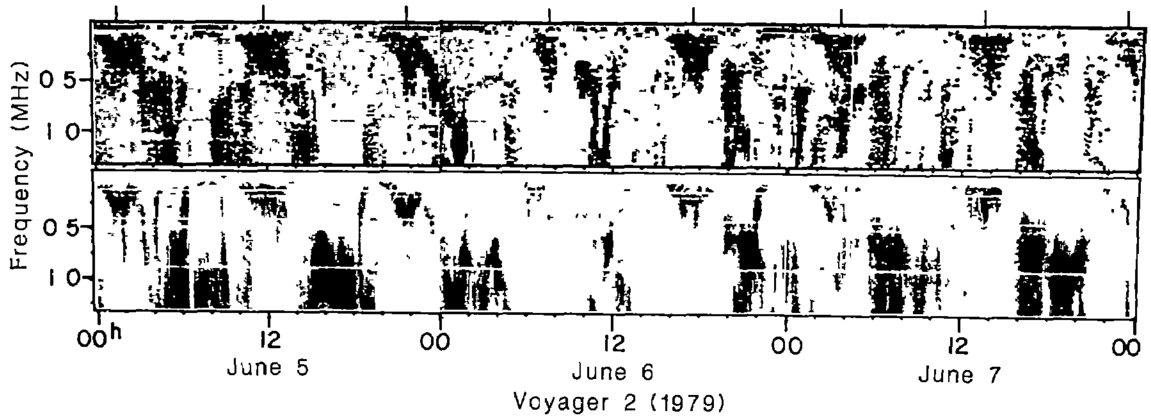


Fig. 1 Voyager 2 spectrogram[5] of jovian radio emissions in the range 20 kHz–1.3 MHz. Total power is shown in bottom panel, increasing darkness being proportional to increasing intensity. Upper panel shows polarization information on the same emissions. In the 72-h period shown, eight consecutive bKOM events are visible as the V shaped emissions extending up to ~500–600 kHz and coincident with the north dipole tip passage: their polarization is right-handed[6]. The events are anticoincident with the hectometric radiation seen at higher frequencies.

the magnetic field line[4], $f_c = 0.028\ B$ kHz being the electron cyclotron frequency for a magnetic field strength of B nT. Thus for a source at the equatorial plane in a dipole magnetic field, the radiation is beamed at angles $\alpha = (\pi/2 - \beta)$ to the equatorial plane, one beam to the north and one to the south. If the emitting region is distributed about the equatorial plane by ~±2° as would be the case if the jovian ESUH waves are the source, then the southern edge of the northern beam would be expected to derive from the most southerly source, that is from −2° magnetic latitude. Typical ray paths derived from ray-tracing calculations in a realistic Io torus model which clearly illustrate this configuration have been presented in Fig. 2 of ref. 10, a simplified sketch is included here in Fig. 2. Had Voyager's excursions in latitude about the equatorial plane been greater, then, according to the theory, they would have observed the high-latitude edges of the beams which, in the case of the northern hemisphere beam, would have derived from the most northerly source, that is from +2° magnetic latitude (see Fig. 2). This effect, where Voyagers appeared to pass right through the beams, has been observed on both Voyager 1 and 2 when outbound at 04.00 and 03.00 LT respectively[6] despite their being at lower jovicentric latitudes (~5°) than when inbound, indicating that at these local times the beams make smaller angles to the magnetic equatorial plane and have a smaller width, possibly due to a smaller source dimension. These sources of bKOM observed by Voyagers when outbound are under study and will be reported separately.

The theory predicts emission of bKOM at the plasma frequency f_p at angles $\beta = \arctan(f_p/f_c)^{1/2}$ with respect to the magnetic field line. Thus assuming a source position at radial distance R_S and magnetic latitude λ_{ms}, the magnetic field magnitude and direction at that point can be computed[11]. Furthermore, assuming a value for f_p within the range of bKOM frequencies observed by Voyager, the theoretical direction of the beam at this frequency can also be computed. If the beam does not pass through the position at which the spacecraft lies when that particular frequency is observed, then the assumed source location is incorrect and another value of R_S is tried until one iterates onto a location that fits. If this technique is applied for all the bKOM frequencies using the corresponding positions of Voyager when they are observed, a radial profile of f_p and therefore of N_e in the Io torus can be constructed. This technique has been applied to terrestrial[12–15] and saturnian[16] myriametric radiations where it could be assumed that the magnetic fields are dipolar, and the source positions could be determined analytically. Jupiter's magnetic field near the equatorial plane at the Io torus and beyond is complicated by the presence of the equatorial current sheet[11] and hence the need for numerical iteration to locate the sources. Voyager 2 inbound data are used because the spacecraft's highest latitude resulted in excursions in and out of the beam during a relatively large number of consecutive jovian rotations. Results from Voyager 1 inbound and Voyagers 1 and 2 outbound will form the basis of a study whose results will be reported separately.

It is assumed, as stated above, that the southern edge of the northern bKOM beam derives from a source at $\lambda_{ms} = -2°$. The positions of Voyager 2 are determined when the leading edge of bKOM at different frequencies is observed. The longitude of Voyager 2 at this time is <202° CML. This is measured for the nine consecutive jovian rotations when the spacecraft was inbound from 95 R_J to 55 R_J. The Io torus plasma density profiles obtained by the remote-sensing technique described above are shown in Fig. 3. Also shown, for comparison, are the profiles determined *in situ* by two experiments[7,17] on Voyager 1 which

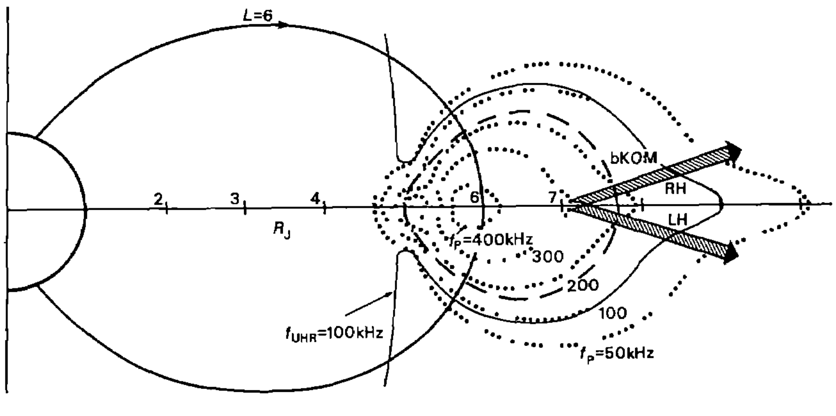


Fig. 2 Schematic of bKOM beaming from the Io plasma torus. Dotted lines represent isoionic contours labelled by plasma frequency f_p. Dotted-dashed line is approximate $f_p = 200$ kHz contour determined *in situ*[1] by Voyager 1. Solid line depicts where upper hybrid resonance frequency $f_{UHR} = 100$ kHz.

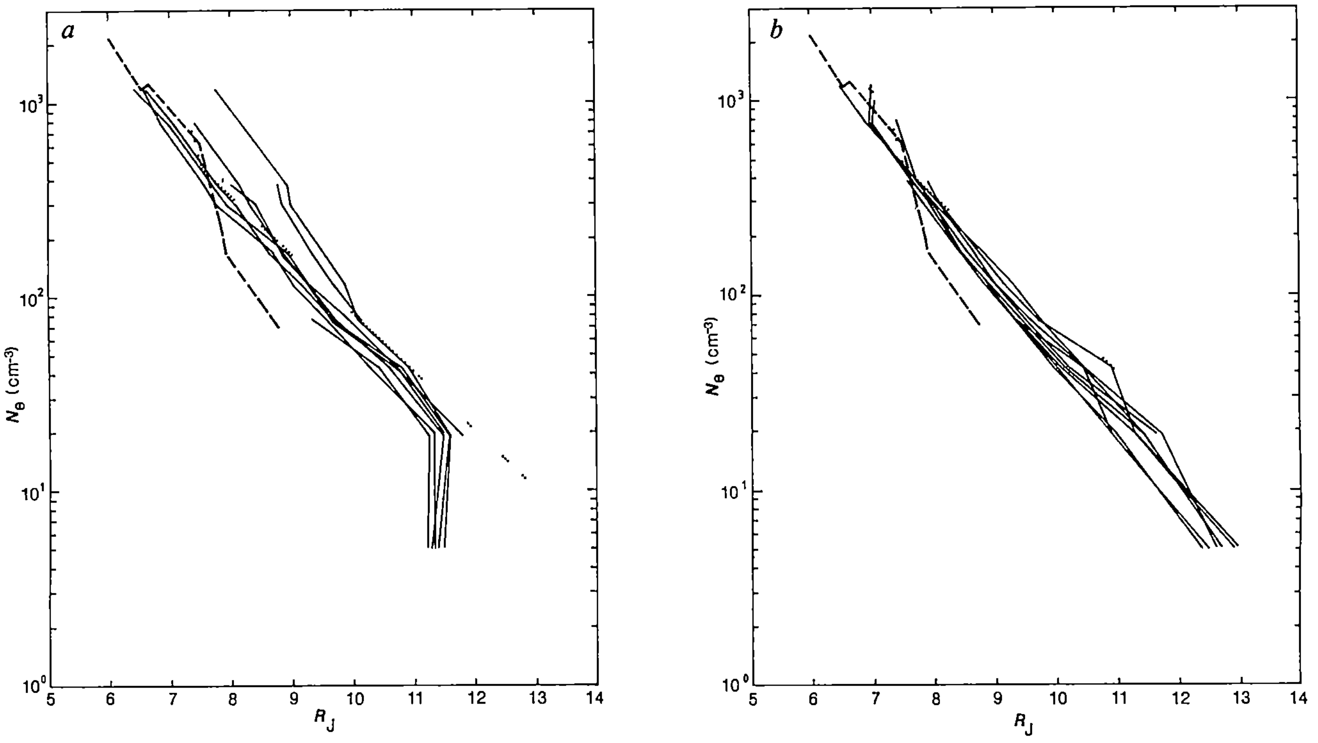


Fig. 3 *a*, Plasma density profiles of the Io torus obtained by remote sensing using the start times of bKOM observed during nine jovian rotations when Voyager 2 was inbound from 95 R_J to 55 R_J. All sources are assumed to be at magnetic latitude $\lambda_{ms} = -2°$. The dashed line represents the profile determined from *in situ* Voyager 1 measurements by the plasma science PLS experiment[17] and the dotted line is derived from measurements of the upper hybrid frequency by the planetary radio astronomy PRA and plasma wave science PWS experiments[7]. *b*, As for *a*, but using the end times of bKOM and assuming all sources are at magnetic latitude $\lambda_{ms} = -1°$.

had traversed the torus 4 months previously. The remotely-sensed profiles are remarkably consistent within themselves considering that there were temporal variations in the torus over the 5 days involved, during which Io made nearly three rotations around Jupiter. The anomalous results at the lowest frequency are possibly due to a wave refraction effect between the source and spacecraft and this is under investigation. The agreement with the Voyager 1 *in situ* profiles is also remarkable considering the time and CML differences between the two sets of measurements. The same remote-sensing exercise has been performed using the trailing edge of the bKOM emissions, again observed in the northern hemisphere but with Voyager 2 moving down in latitude towards the equatorial plane. Figure 4 shows the profiles obtained for a source latitude of $\lambda_{ms} = -1°$ which was found to give better agreement with the *in situ* profiles than $\lambda_{ms} = -2°$. This difference may be due to small differences in the widths or positions of the source regions with CML since these results correspond to longitudes >202° CML, no *in situ* results exist for the variation with CML of the latitudinal width of the source ESUH waves, although certain asymmetries with respect to the equatorial plane have been observed[18]. The difference may also arise from the fact that the centrifugal equator, which is the symmetry plane for the Io torus plasma, has a magnetic latitude which differs by ~1° between the two cases due to their CML differences[19].

Note that the slopes of the remotely-sensed profiles differ only slightly from that of the *in situ* profile determined from the observations of upper hybrid waves (dotted line). If one had assumed that the source latitude decreased by $\simeq 0.5°$ with increasing distance from 8 R_J to 13 R_J the slopes would be identical. Such a decrease in the width of the ESUH emission regions with distance has been observed[18].

Because the beam widths depend to a first approximation on the latitudinal dimensions of the source, further computations show that for Voyager 2 to have observed the top edge of the northern beam when inbound on the dayside, the source ESUH waves would have had to be restricted to within $\sim\lambda_{ms} \simeq 0.5°$ of the equatorial plane and even then the effect would only have been visible at the higher frequencies. Alternatively, had Voyager 2's excursion into the northern hemisphere been $>\lambda_m \simeq 20$-$25°$ it would have observed the whole beam at all frequencies.

In conclusion, bearing in mind the temporal and longitudinal variations of the torus plasma density and of the current sheet magnetic field which has been assumed constant in the model used, it is difficult to imagine that the remotely-sensed profiles obtained using data from a spacecraft $>3\times10^6$ km from the source could be so consistently similar to the *in situ* profiles unless the linear theory proposed is valid.

I thank Yolande Leblanc, Observatoire de Paris, for providing the PRA data and for much advice and information, Jack Connerney, NASA Goddard Space Flight Center, for supplying the magnetic field program and Fran Bagenal, Imperial College, London, for helpful discussions.

Received 27 May, accepted 4 August 1986.

1. Warwick, J. W. *et al. Science* **204**, 995-998 (1979).
2. Warwick, J. W. *et al. Science* **206**, 991-995 (1979).
3. Kurth, W. S., Gurnett, D. A. & Scarf, F. L. *Geophys. Res. Lett.* **7**, 61-64 (1980).
4. Jones, D. *Nature* **288**, 225-229 (1980).
5. Desch, M. D. & Kaiser, M. L. *J. geophys. Res.* **85**, 4248-4256 (1980).
6. Leblanc, Y. & Daigne, G. *J. geophys. Res.* **90**, 12073-12080 (1985).
7. Barbosa, D. D. & Kurth, W. S. *J. geophys. Res.* **85**, 6729-6742 (1980).
8. Oya, H. *Planet. Space Sci.* **22**, 687-708 (1974).
9. Jones, D. *Nature* **260**, 686-689 (1976).
10. Jones, D. *Adv. Space Res.* **1**, 333-336 (1981).
11. Acuna, M. H., Behannon, K. W. & Connerney, J. E. P. *Physics of the Jovian Magnetosphere* (ed. Dessler, A. J.) 1-50 (Cambridge University Press, 1983).
12. Jones, D. *Nature* **294**, 728-730 (1981).
13. Jones, D. *Planet. Space Sci.* **30**, 399-410 (1982).
14. Jones, D. *J. Geophys.* **52**, 158-166 (1983).
15. Budden, K. G. & Jones, D. *Comparative Study of Magnetospheric Systems* (ed. CNES) 563-580 (Cepaudes-Editions, Toulouse, 1986).
16. Jones, D. *Nature* **306**, 453-456 (1983).
17. Bagenal, F. & Sullivan, J. D. *J. geophys. Res.* **86**, 8447-8466 (1981).
18. Kurth, W. S., Barbosa, D. D., Gurnett, D. A. & Scarf, F. L. *Geophys. Res. Lett.* **7**, 57-60 (1980).
19. Dessler, A. J. *Physics of the Jovian Magnetosphere* (ed. Dessler, A. J.) 498-504 (Cambridge University Press, 1983).

An icy-glue model of cometary nuclei

Tamas I. Gombosi* & Harry L. F. Houpis*

Central Research Institute for Physics, Hungarian Academy of Sciences, H-1525 Budapest 114, POB 49, Hungary

Since 1950 a number of models have been proposed to explain the observations of comets. For the most part, the icy conglomerate model of Whipple[1] has been the standard which others have followed. In light of the recent images made of comet Halley, by the Vega and Giotto spacecraft, we suggest here that the nucleus is composed of rather large (tens of centimetres to hundreds of metres) porous refractory boulders, 'cemented' together with an ice–dust grain mix of the icy conglomerate type. This model retains the advantages of the icy conglomerate formalism on a local level, and also introduces a new global framework. We discuss the model in terms of the improving knowledge about comets, the creation of the nucleus in the cosmogonic sense, and expectations of the future analysis of data and observations.

As well as explaining the production rate of various chemical species in the cometary coma, previous nucleus modelling has been constrained by the following observational facts: the light-curve of a comet is asymmetric with respect to perihelion; most short- and intermediate-period comets ($P < 200$ yr) 'turn on' inside ~ 3 AU; some new comets show gas activity at very large heliocentric distances (5–10 AU); and comets have been known to split. In the past few years, a number of important observations have been made which further constrain a nucleus model. Radar observations of comet IRAS–Araki–Alcock[2] have shown that the nucleus is highly non-spherical, that the surface is rough on a scale larger than the radar wavelength (3.54 and 12.9 cm), and that there is an expanding 'skirt' of non-gravitationally bound material with individual radii larger than a few centimetres. The clumpy nature of OH emissions in the atmosphere of comet Halley[3] is suggestive of extended sublimation sources.

These most recent observations, together with an assortment of other evidence, led Weissmann[4] to postulate a primordial rubble pile of small icy conglomerates. Although this model, and similarly that of Donn *et al.*[5], can also explain the splitting of comets as the 'breaking free of weakly bonded fragments', the recent images of the comet Halley nucleus by the Vega and Giotto spacecraft[6–8] suggest another way of looking at this process as well as the nucleus as a whole. The shape of the Halley nucleus is quite non-spherical ($\sim 7.5 \times 7.5 \times 15\ \text{km}^3$) with altitude variations of ~ 100 m. About seven major jets, some of them grouped into bundles (the source diameter of each bundle being ~ 3 km, whereas that of the single dust jets appears to be ~ 500 m) indicate an origin on the sunlit side of the nucleus and point in the general direction of the Sun. The source area of all the active jets is estimated to be $\sim 10\%$ of the total surface area. The rest of the surface appears to be inactive, although a low, undetected dust activity cannot be dismissed. Furthermore, the surface has a rather uniform albedo of a few per cent, indicative of a porous dust coating with little or no ice; this is also consistent with the infrared measurements of an average surface temperature over 300 K.

These latest results from the spacecraft encounters with comet Halley lead us to propose a new model of the nucleus. We suggest that the nucleus is composed of rather large (tens of centimetres to hundreds of metres) porous refractory boulders (similar in composition to the outer asteroid belt objects and the irregular outer satellites of the giant planets; that is, hydrous silicates containing complex carbon compounds) 'cemented' together with an ice–dust grain mix of the icy conglomerate type ('Whipple glue'). As the nucleus evolves with repeated passages close to the Sun, those regions above the largest boulders, which may be only a few metres thick, become depleted in their volatile ices and hence inactive. We call these regions dust plates, and they may be only a few centimetres thick. On the other hand, the regions between the largest boulders, where there are the smaller boulders and the ice–dust mix, may penetrate quite deep into the nucleus (50–100 m) and hence are not likely to become inactive for several (perhaps 50–100) orbits around the Sun. We call these regions macropores to distinguish them from the micropores in the Whipple glue. It is from here that the jets originate. Figure 1 shows conceptual drawing of such a nucleus.

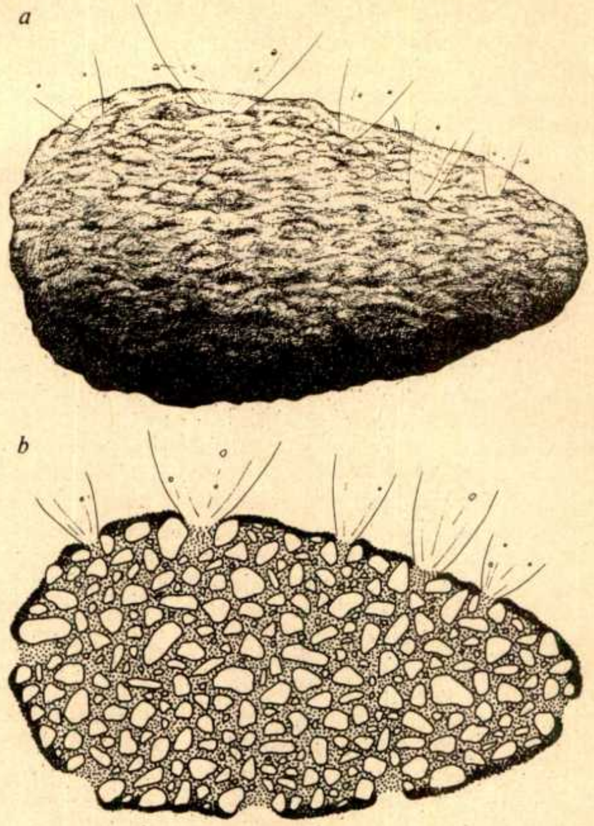


Fig. 1 *a*, Three-dimensional perspective of a comet-Halley-shaped icy-glue cometary nucleus showing active and inactive macropores, release of small boulders, surface altitude variations of porous refractory framework imbedded in Whipple glue, outer dust covering, and penetrations of the form surface by potentially active macropores.

This model explains in a natural way the above observations. The general shape of the nucleus and its surface altitude variations are a direct consequence of the low self-gravity and the large boulder framework, whereas the uniform porous dust covering is a result of the volatile ices receding below the surface and the creation of dust plates.

Jets, as the major sources of cometary activity, are the result of a group of macropores coming into view of the Sun. Depending on the placement of the boulders, the jets may appear as line sources delineating the shape of the large boulders near the surface or as point sources for a more evolved surface topology. These appearances are reminiscent of the surface regions sketched by Sekanina and Larson[9,10]. Furthermore, the latest analysis of the Vega images indicates the existence of both linear and point sources on comet Halley (B. A. Smith, personal communication).

The percentage of comet Halley's surface capable of significant dust emission can be estimated from the observed jet activity. Using a total active area of $\sim 50\ \text{km}^2$ as seen by Keller *et al.*[6] for the Halley nucleus, one can conclude that as much as 44% of the entire surface is capable of jet activity.

Typically, we expect the width of the smallest dimension of a jet at the surface to be no more than a few hundred metres. However, a few hundred metres from the surface the jet will have expanded laterally due to the existence of a gas-pressure gradient between the site above the macropore and that of the surrounding dust plates[11].

---

* Present address: Space Physics Research Laboratory, Space Research Building, University of Michigan, Ann Arbor, Michigan 48109, USA (T.I.G.); Centre of Astrophysics and Space Science, University of California, San Diego, California 92093, USA (H.L.F.H.).

The literature often suggests that jets may originate from fissures due to thermal stress, which have opened on the surface. The freshly exposed ice then provides the source for the jets. Our contention is that although the surface may crack, the resulting excess production will be of an outburst nature rather than a steady jet. The reason is that exposed ice will quickly recede to form a dust mantle (only a few tenths of a centimetre are needed) and return the excessively active surface to that of a diffusion process. Thus, although fissures and unglued small boulders may account for outbursts, the macropores are responsible for the steady jet activity.

Splitting of comets is now explained by pieces of the framework becoming unglued; that is, the Whipple glue around a particular boulder has been sublimed and entrained to such a point, that the boulder eventually breaks free. This process which is obviously random with respect to the comet's orbital parameters and heliocentric distance, explains why more splitting fragments are a small fraction of the original nucleus, separating with a small relative velocity and disappearing visually in ~100 days, and allows for 'flaring' activity as the new uncovered Whipple glue is exposed to the solar environment. The decrease in splitting frequency as a comet spends more time in the inner Solar System may be due to such processes as densification of the ice-dust grain mix[12], and the recondensation near the surface of volatile ices which have sublimed from the comet interior[13]. We also suggest that outbursts from comets may be in part the result of the much smaller boulders becoming unglued. We note that a major difference between our mechanism and that of the rubble-pile model is that splitting of the icy glue nucleus is a natural consequence of the structural make-up, whereas the rubble-pile theory must invoke 'weakly bonded fragments' and the necessity for a heat wave to have some *ad hoc* preference to propagate along the interfaces between them.

The activity of some 'new' comets at large heliocentric distance may also be partially explained by postulating that these comets are indeed new to the inner Solar System, with the entire surface capable of dust and ice production, and that comets are probably significantly larger than previously thought. Thus, it is not until the inactive dust plates begin to predominate after a few orbits that activity at large heliocentric distances begins to subside. This idea may of course be used in conjunction with the chemical differentiation model[14] to account for the 'tapering off' of activity at smaller heliocentric distances. The asymmetric behaviour of cometary light curves about perihelion can be explained in part by the location of the active macropores.

For the creation of an icy-glue nucleus we suggest the following model. The large porous boulders first agglomerated in the pre-Jupiter region. The growing Jupiter and Saturn protoplanets caused the diffusion of a substantial amount of this material through the pre-Uranus-Neptune zone. Because agglomeration of the original influx of material into this outer zone was quite slow, as suggested by a number of theories[15], the boulders arrived to find the ice and dust still in the small-grain phase and somewhat extended about the ecliptic plane. The boulders and grains then came together to form the icy-glue nuclei described previously, with the Uranus and Neptune protoplanets ejecting them to the Oort cloud.

Although this model is only illustrative and needs to be subjected to a proper dynamical analysis, we point out an important difference between the icy-glue complex and the icy conglomerate models. On the one hand, the icy conglomerate model has its basic building blocks constructed from both dust and ice grains, although the mixture may be heterogeneous relative to the centre of the nucleus or the individual building blocks. On the other hand the icy-glue model is built from two different types of basic units: porous refractory boulders with little or no initial volatile content, and grains of dust and ice. This fundamental difference may prove useful in constructing a cosmogonic theory of the Solar System.

We thank Drs Bradford Smith, Akos Körösmezey, Steven M. Larson and Károly Szegõ for useful discussions, Carole Turner for the careful editing of this manuscript and Diane McGregor for the artwork of Fig. 1. Part of this work was supported by NSF grant AST-8605994.

Received 15 May; accepted 11 August 1986.

1. Whipple, F. L. *Astrophys J.* **111,** 375-349 (1950).
2. Goldstein, R. M., Jurgens, R. F. & Sekanina, Z. *Astr. J.* **89,** 1745-1754 (1984).
3. dePater, I., P. & Schenewerk, M. *Pap. presented at Conf. Am. Geophys. Un.*, San Francisco, December 1985; *IAU Circ.* No. 4142 (1985).
4. Weissman, P. *Nature* **320,** 242-244 (1986).
5. Donn, B., Hughes, D. W. & Daniels, P. A. *Bull. Am. astr. Soc.* **17,** 520 (1985).
6. Keller, U. *Pap. presented at Max Planck Institut für Aeronomie Seminar*, Katlenburg-Lindau, W. Germany, 2 April 1986; Keller U. *et al Nature* **321,** 320-326 (1986).
7. Sagdeev, R. Z. *et al. Nature* **321,** 259-262 (1986).
8. Combes, M. *et al. Nature* **321,** 266-268 (1986).
9. Sekanina, Z. *Astr. J.* **86,** 1741-1773 (1981)
10. Sekanina, Z. & Larson S. *Nature* **321,** 357-361 (1986) *Astr. J.* (submitted).
11. Kitamura, Y. *Icarus* (submitted).
12. Smoluchowski, R., Marie M. & McWilliam, A. *Earth, Moon & Planets* **30,** 281-288 (1984).
13. Whipple, F. L. & Stefanik, R. P. in *Nature et Origine des Comètes*, 13th Liège Symp., 33-51 (1966).
14. Houpis, H. L. F. Ip, W.-H., & Mendis, D. A. *Astrophys. J.* **245,** 654-667 (1985).
15. *Protostars and Planets II* (eds Black, D. C. & Matthews, M. S.) Part VII, pp. 959-1155 (University of Arizona Press, 1985).

Distribution of energetic particles near interplanetary shocks

A. Hewish* & S. Bravo†

* Mullard Radio Astronomy Observatory, Cavendish Laboratory,
Madingley Road, Cambridge CB3 0HE, UK
† Instituto de Geofisica, Ciudad Universitaria, Codigo 04510,
Mexico D.F.

The acceleration of charged particles by collisionless shocks is a process of considerable importance in astrophysics, but the physical mechanism is not yet understood in detail. Energetic particles near interplanetary shocks may be studied *in situ* by spacecraft, but little is known about their distribution across the shock surface because of inadequate knowledge of the location of the sampling point in relation to the whole disturbance. A ground-based method has been developed[2] which enables shock-associated disturbances to be mapped in three dimensions, and we have used it, in combination with near-Earth spacecraft measurements, to study how the particle flux varies for shocks travelling at different angles to the Sun-Earth line. We have observed a pronounced east-west asymmetry which indicates that the inclination of the interplanetary magnetic field to the shock surface must be a potent factor in the acceleration process.

The method of mapping the disturbances was based on daily observations of a large grid of radio sources, the scintillation of which was related to the mean plasma density along many different lines of sight[1]. Examples of the structures of large-scale disturbances revealed by this method have been described previously[2-4]. It was found that strong shocks typically occurred at the front of enhanced density shells which subtended a solid angle of $\sim\pi/2$ sr at the Sun. The shells were compression zones, caused by sudden eruptions of fast solar wind which lasted for several days, and they could always be traced back to sources coincident with coronal holes[5]. Among the 96 disturbances detected during August 1978-September 1979, 73 were classified as 'erupting streams'; the remainder were co-rotating interaction regions of longer duration[6]. All significant shocks (shock strength $\geqslant 1.5$) recorded by near-Earth spacecraft were accompanied by compression zones that were mapped by our method.

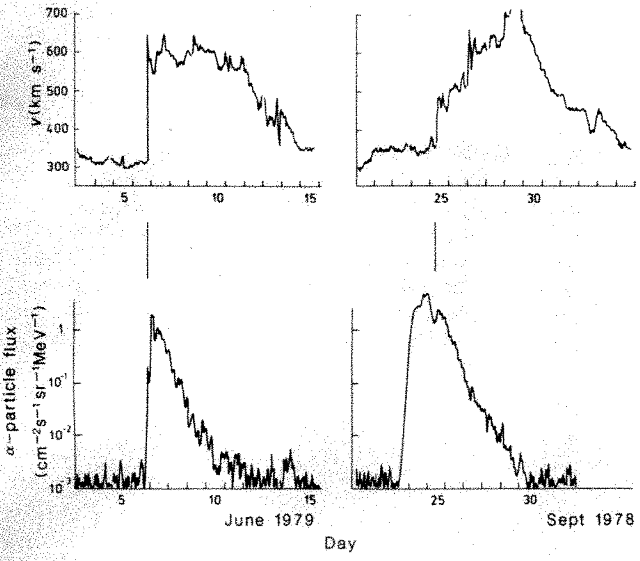


Fig. 1 Temporal variations of solar wind speed (v) and α-particle flux in the range 9–70 MeV associated with two strong interplanetary shocks. Vertical line, time of shock passage at the Earth.

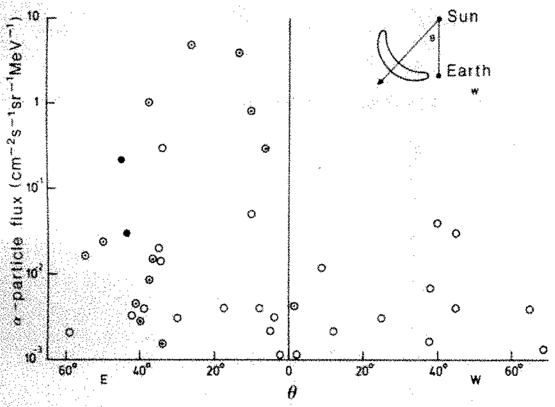


Fig. 2 Peak α-particle flux observed for shocks travelling at different angles, θ, to the Sun–Earth line. Shock strengths: ●, ⩾3; ⊙, ⩾2; ○, <2.

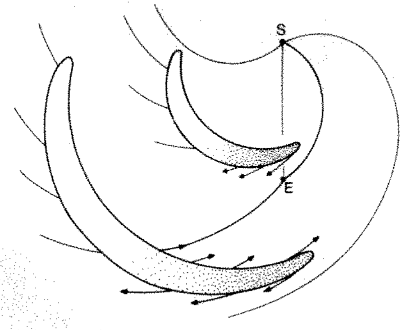


Fig. 3 Schematic diagram showing the regions of greatest particle acceleration when the shock disturbance has reached distances from the Sun (S) 1 and 2 AU. The spiral magnetic field lines correspond to a solar wind speed of 350 km s^{-1}. Arrows denote the directions of particles moving away from the shock. The field line through the Earth (E) beyond 1 AU is calculated for a speed of 600 km s^{-1}, appropriate to the high-speed flow behind the shock.

However, we sometimes detected eruptions unaccompanied by shocks, either because the shocks were too weak or because they did not strike the Earth.

All the disturbances that were classed as erupting streams, and that also caused sudden geomagnetic storms, are listed in Table 1 together with spacecraft measurements of the shock strength. Also given are the estimated directions of travel (θ), relative to the Sun–Earth line, obtained by our method. As described previously, the angles were obtained by comparison with theoretical maps computed for simple models[6]. The accuracy of ($\sim\pm20°$) is not high, but there was usually no difficulty in placing disturbances to the east or west of the Sun–Earth line.

For the energetic particle data we used the α-particle flux in the range 9–70 MeV from the GMS/SEM satellites[7]. During the largest particle events the flux increased by a factor of $\sim10^3$ (Fig. 1). Table 1 gives the peak flux associated with each shock, smoothed over a few hours. When the particle fluxes are plotted against the directions of travel (Fig. 2), a pronounced asymmetry is evident. Only shock disturbances centred on directions to the east of the Sun–Earth line produced the largest enhancements of flux. The background flux was typically $\sim2\times10^{-3}$ cm^{-2} s^{-1} sr^{-1} MeV^{-1}, and enhancements greater than four times this value occurred on average only once or twice per month. At this level there were 16 events associated with shocks passing on the east, compared with 3 in the west. Above 5×10^{-2} cm^{-2} s^{-1} sr^{-1} MeV^{-1}, there were 8 shocks in the east and none in the west.

Table 1 Shock disturbances striking the Earth

Date		Arrival time UT	θ	Particle flux	Shock strength
1978	18 Aug	1242	38 W	1.5×10^{-3}	1.4
	27 Aug	0246	43 E	4×10^{-3}	2.5
	5 Sep	1859	5 E	2×10^{-3}	1.6
	9 Sep	0354	38 W	7×10^{-3}	—
	11 Sep	0900	69 W	1.5×10^{-3}	—
	20 Sep	1417	30 E	3×10^{-3}	—
	25 Sep	0718	26 E	5.0	>2
	29 Sep	0301	9 W	1.2×10^{-2}	1.8
	4 Oct	0047	37 E	1.3×10^{-2}	2.2
	9 Oct	0332	35 E	2×10^{-2}	1.5
	17 Oct	0430	4 E	3×10^{-3}	1.4
	29 Oct	1116	42 E	3.5×10^{-3}	1.7
	8 Nov	0152	59 E	2×10^{-3}	1.5
	12 Nov	0100	6 E	3×10^{-1}	2.8
	18 Nov	2320	2 E	1×10^{-3}	1.2
	18 Dec	0036	12 W	2×10^{-3}	—
	25 Dec	1212	40 E	2.8×10^{-3}	2.0
1979	4 Jan	1725	45 W	4×10^{-3}	—
	6 Jan	0032	40 W	4×10^{-2}	1.7
	9 Jan	0340	55 E	1.5×10^{-2}	2.2
	25 Jan	0139	39 E	3.8×10^{-3}	1.9
	21 Feb	0302	50 E	1.3×10^{-2}	2.2
	4 Mar	0445	35 E	1.5×10^{-2}	—
	6 Mar	0818	45 W	3×10^{-2}	1.5
	9 Mar	0808	25 W	3×10^{-3}	1.7
	17 Mar	0231	65 W	4×10^{-3}	—
	22 Mar	0826	34 E	1.3×10^{-3}	2.5
	5 Apr	0156	37 E	1.0	2.0
	24 Apr	2358	43 E	3×10^{-2}	3.3
	29 May	1851	38 E	8×10^{-3}	2.2
	6 Jun	1927	13 E	4.0	2.9
	6 Jul	1930	45 E	2×10^{-1}	3.2
	12 Jul	1240	7 E	3×10^{-3}	—
	11 Aug	1812	18 E	4×10^{-3}	1.6
	13 Aug	0639	2 W	1×10^{-3}	—
	20 Aug	0625	10 E	8×10^{-1}	2.4
	24 Aug	1457	10 E	5×10^{-2}	1.7
	29 Aug	0459	34 E	3×10^{-1}	—

The shock strengths (downstream/upstream density) were taken from refs 13, 14. The α-particle flux is in units of cm^{-2} s^{-1} sr^{-1} MeV^{-1}.

Assuming that all the shocks are similar, it follows that the particle flux at the shock surface follows the distribution shown in Fig. 2, but in the opposite sense. Thus, a shock travelling directly towards the Earth would generate a flux which peaked at 20-30° west of the Sun-Earth line.

The asymmetry can be understood if the acceleration depends on the angle between the magnetic field and the shock surface. Figure 3 shows the geometry for an idealized interplanetary field, assuming an ambient solar wind of 350 km s^{-1}. For co-rotating shocks beyond 1 AU caused by a stable high-speed stream, the maximum particle flux occurs when the shock surface is nearly parallel to the field[8] (the 'quasi-perpendicular' case). For a spherical shock front appropriate to an erupting stream, it is clear from Fig. 3 that this condition can occur on one side of the disturbance only, and the asymmetry is in the correct sense to explain our observations. Another factor favouring the parallel field condition is the co-rotation of the stream following its sudden eruption. This will tend to increase the curvature of the shock front on the same side of the disturbance.

In addition to the asymmetry, the characteristic sharp rise and slow decay of the flux (see Fig. 1) is readily explained. Recent evidence[9,10] supports a model in which particles are accelerated continuously near the shock and then diffuse away in both directions along the field lines. From the geometry in Fig. 3 it may be seen that for a shock passing to the east, the field line through the Earth makes connection only shortly before the shock arrives, but thereafter remains connected for several days as the disturbance propagates beyond 1 AU. The systematic eastward movement of the point of contact as the shock recedes from the Earth effectively scans the particle distribution in a different way.

In previous studies of energetic particle events, associations with solar flares have been assumed and conflicting results have been obtained. It has been deduced, for example, that flares are most effective when they occur on the west of the disk, which supports the theory that energetic particles are released at the flare site and channeled along 'well-connected' field lines to the Earth[11]. By contrast, more recent work has suggested that the flux is greatest for shocks associated with flares on the east of the disk[12]. These discrepancies are not surprising in the light of our observations that shocks are caused by eruptions generated by coronal holes and not by flares[6]. Our present conclusions are qualitatively similar to those of ref. 12, but differ in showing a larger and more sharply defined east-west asymmetry for the particle flux.

Received 23 June; accepted 24 July 1986.

1. Tappin, S. J. *Planet. Space Sci.* **34,** 93-97 (1986).
2. Gapper, G. R., Hewish, A., Purvis, A. & Duffett-Smith, P. J. *Nature* **296,** 633-636 (1982).
3. Tappin, S. J., Hewish, A. & Gapper, G. R. *Planet. Space Sci.* **31,** 1171-1176 (1983).
4. Tappin, S. J., Hewish, A. & Gapper, G. R. *Planet. Space Sci.* **32,** 1273-1281 (1984).
5. Hewish, A., Tappin, S. J. & Gapper, G. R. *Nature* **314,** 137-140 (1985).
6. Hewish, A. & Bravo, S. *Sol. Phys.* (in the press).
7. Nagai, T. & Kohno, T. *Sol. terr. envir. Res. Japan* **8,** 20-33 (1984).
8. Tsurutani, B. T., Smith, E. J., Pyle, K. R. & Simpson, J. A. *J. geophys. Res.* **87,** 7389-7404 (1982).
9. Evenson, P., Meyer, P. & Yanagita, S. *J. geophys. Res.* **87,** 625-631 (1982).
10. Sanahuja, B., Domingo, V., Wenzel, K.-P., Joselyn, J. A. & Keppler, E. *Sol. Phys.* **84,** 321-337 (1983).
11. Van Hollebeke, M. A. I., Ma Sung, L. S. & McDonald, F. B. *Sol. Phys.* **41,** 189-193 (1975).
12. Sarris, E. T., Decker, R. B. & Krimigis, S. M. *J. geophys. Res.* **90,** 3961-3965 (1985).
13. van Nes, P., Reinhard, R., Sanderson, T. R. & Wenzel, K.-P. *J. geophys. Res.* **89,** 2122-2132 (1984).
14. Cane, H. V. *J. geophys. Res.* **90,** 191-197 (1985).

Searching potential energy surfaces by simulated annealing

L. T. Wille

SERC Daresbury Laboratory, Daresbury, Warrington WA4 4AD, UK

Many problems in physics[1], chemistry, biology[2] and mathematics[3] involve the determination of the absolute minimum of a certain multidimensional function. In most cases of practical interest this is a complicated matter, owing to the presence of local minima, and even more so because the number of local minima often increases exponentially with the problem size. Standard techniques apply local optimizers to many random initial configurations, but soon become intractable as the dimensionality increases. Here I show how the simulated annealing method can be used to guide a search towards the absolute minimum. The method is illustrated on a problem recently discussed in this journal, namely the minimum-energy configuration of equal charges confined to a sphere. This problem, although easy to visualize, can be used to simulate much of the complexity of the above problems by considering a large number of particles. In this limit several minima have been found that have previously been missed by other authors using classical techniques.

To determine the ground-state configuration of a system of interacting particles one needs to find the absolute minimum on the potential-energy surface (strictly speaking it is the free energy that has to be minimized, but one usually assumes the system to be at zero temperature). A local minimum can be found by one of many possible search techniques[4], most of which are based on a gradient method. Starting from a certain point on the potential energy surface one takes a step in a direction determined by the line of steepest descent and repeats this process until a minimum is reached, to within a certain accuracy. The collection of all points from which the same (local) minimum is obtained by this algorithm is called the catchment region of the minimum. This whole procedure is repeated a number of times starting from random initial configurations and produces a list of local minima, the smallest of which is taken to be the global minimum. Obviously there is no guarantee that the global minimum will be found in this way, especially if the catchment region is narrow. Alternatively, instead of taking random initial configurations one tries to deduce growth schemes from low-dimensional systems. The algorithm is then applied to configurations obtained from a set of growth rules. This procedure often produces highly symmetric configurations; the previous remarks are then even more pertinent.

Obviously a method that avoids getting trapped in local minima would increase the likelihood of detecting the global minimum. It has been shown[5,6] that a continuous version of the simulated annealing method proposed by Kirkpatrick *et al.*[7] for discrete problems can be used for this purpose. An overview of simulated annealing and its applications has been presented elsewhere[8] and only a brief description will be given here. Consider the potential energy to be minimized as $E(\mathbf{r}_1, \ldots, \mathbf{r}_N)$, where the $\mathbf{r}_i$, $i = 1, \ldots, N$ are the positions of the particles. The simulated annealing method uses the Metropolis prescription[9] to decide whether or not to accept a random step $\Delta\mathbf{r}_i$, $i = 1, \ldots, N$. If the associated energy change $\Delta E = E(\mathbf{r}_1, \ldots \mathbf{r}_i + \Delta\mathbf{r}_i, \ldots, \mathbf{r}_N) - E(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_N)$ is negative, the step is accepted, otherwise it is accepted with a probability $P = \exp(-\Delta E/T)$, where T is the temperature (in reduced units). Obviously this corresponds to a random walk on the potential energy surface, where the walker feels attracted to minima. In the process of annealing the system is started in a random initial configuration at a sufficiently high temperature. Several series of steps are attempted and the temperature is then reduced by a certain factor (typically 0.75-0.9). This process is repeated until no further improvements have been made over a number of iterations. This scheme is similar to the way a crystal is grown from the melt. If the fluid is cooled too quickly (or quenched) a

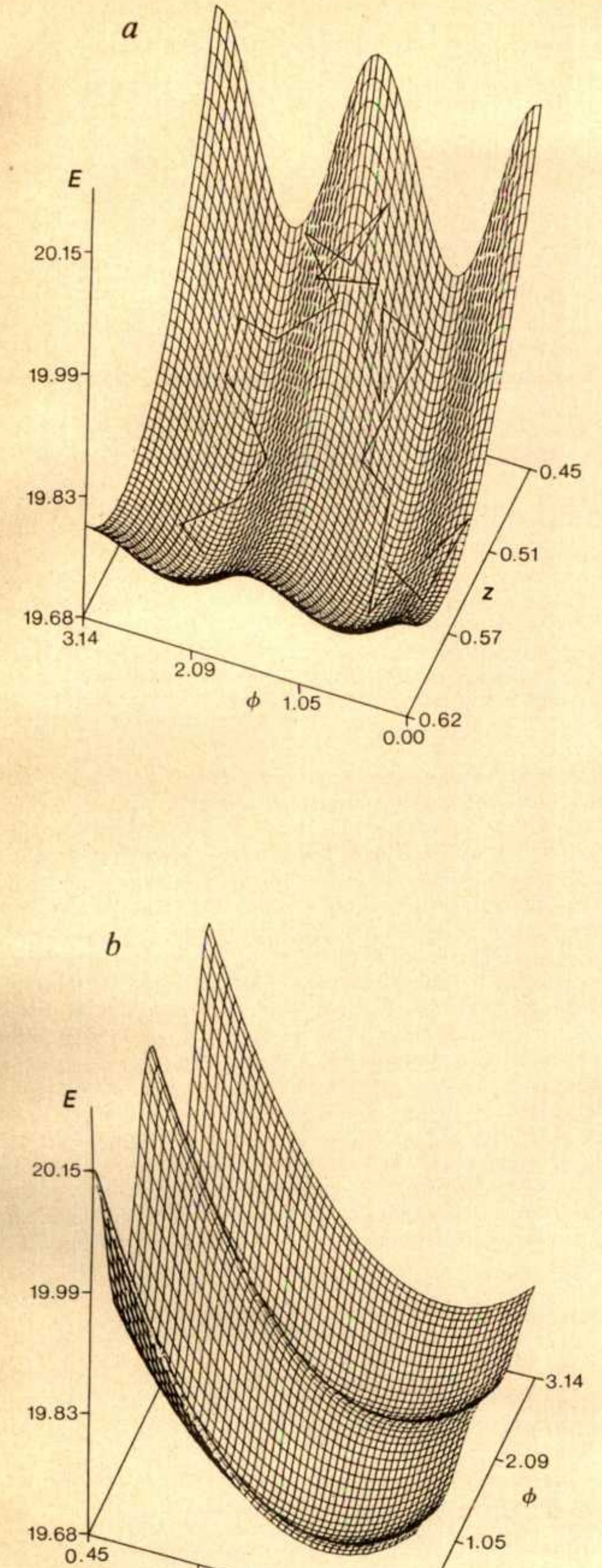


Fig. 1 The potential energy surface for an antiprismal arrangement of charges: z is the distance of the planes from the equator and ϕ is the relative rotation angle. *a*, The random walk towards the minimum. *b*, A cubical arrangement $\phi = \frac{1}{2}\pi$, $z = 0.577$ corresponding to a saddle point on the potential energy surface.

disordered structure with many defects results. This would correspond to a local minimum. On the other hand, the true crystalline ground state can be obtained by a process of slow annealing. It was this powerful analogy which motivated Kirkpatrick *et al.* to develop the simulated annealing method[7].

In the calculations presented here the initial temperature was taken from a short preliminary simulation and the step length was adjusted so that 25–50% of the steps were accepted. Typically 100 moves for each particle were attempted before the temperature was reduced by a factor of 0.9. However, it was found that in many cases a much faster cooling produced the same results. As soon as the energy dropped below a certain value (also taken from a preliminary simulation) a steepest-descent optimizer was called. Finally the whole process was repeated ten times for each number of particles. This precaution was taken because it was found previously[6] that, in spite of slow cooling, trapping in a low-lying local minimum can still occur. The use of a local optimizer in the later stages of the simulation is also an innovation and gives rise to a great reduction in computation time.

Recently there has been renewed interest in the minimum-energy configurations of systems of charged particles confined to a circle or a sphere[10–17]. The problem of the distribution of equal point charges on a sphere is long standing; it was investigated at the beginning of this century by J. J. Thomson and D. Hilbert and later by many mathematicians (for a historical review see ref. 18). Very little analytical progress has been made and most of the later work[19–21] uses numerical methods to find the global minimum for this and other extremal problems over a spherical surface. It has been shown[22] that if the configuration possesses rotational symmetry about the diameter through every particle, this configuration is in equilibrium whatever the force law is, provided that it depends only on the distance between the particles. In this work the charges are assumed to interact via a Coulomb potential $V = 1/r$. It is then known that there are no minimum-energy configurations with charges inside the sphere because the potential is harmonic, that is, it satisfies the Laplace equation[23,24]. Hence the particles can be restricted to move on the surface of the sphere during the simulation, a simplification that would no longer be possible for a potential of the form $1/r^n (n > 1)$.

Table 1 shows the minimum-energy configurations for N unit charges ($5 \leq N \leq 20$) confined to a sphere with unit radius, determined by the simulated annealing method. These configurations have been rotated so that the major symmetry axis lies along the north–south direction. The charges are grouped in planes parallel to the equator. In the planes the charges form polygons which are rotated relative to one another. A notation such as $15^3 1$ (height 0.610) means that there is one charge at the north pole, 5 charges at a height 0.610 (in units of the sphere radius), 5 in the equator plane, 5 at a height −0.610 and one at the south pole. The surprising fact that the charges tend to form parallel polygons has been noticed by previous authors. However, this is not a universal feature as is illustrated by the low-symmetry solution for $N = 11$, missed in ref. 17, but found by other authors[19,21]. Other improvements over the results in ref. 17 for $N = 13$ and 15 have also been given before[21]. However, the present results disagree with published results for $N = 14$, which according to ref. 21 should form a $14^3 1$ configuration (with an energy of 69.496), but which is found here and in ref. 17 to form a $16^2 1$ configuration. Finally, in the range $N \leq$ 17–20 which was investigated for the first time in ref. 17, improvements have been found for $N = 19$ and $N = 20$.

Figure 1 illustrates visually how simulated annealing works, by showing the results of an actual simulation for $N = 8$. It is known that when $N = 8$ the charges do not occupy the corners of a cube, and in general the Platonic solids do not necessarily produce the minimum-energy configuration. Although Leech's analysis[12] shows that charges placed on the corners of a regular polyhedron are in equilibrium, this is not in general a stable

Table 1 Minimum energy configuration for N charges confined to a sphere, interacting via a Coulomb potential

N	Configuration	Heights	Energy
5	1 3 1		6.4747
6	1 4 1		9.9853
7	1 5 1		14.453
8	4^2	0.560	19.675
9	3^3	0.704	25.760
10	1 4^2 1	0.423	32.717
11	1 2 4 2^2	0.515, 0.168, −0.553, −0.806	40.596
12	1 5^2 1	0.447	49.165
13	1 2^2 4 2^2	0.611, 0.501, −0.104, −0.545 −0.863	58.853
14	1 6^2 1	0.455	69.306
15	3^5	0.827, 0.407	80.670
16	4^4	0.785, 0.197	92.920
17	1 5^3 1	0.610	106.05
18	1 4^4 1	0.675, 0.203	120.08
19	1 4 2 4 2^2 4	0.638, 0.540, −0.071, −0.197 −0.333, −0.823	135.09
20	1 3 6^2 3 1	0.693, 0.542	150.88

Results obtained by the simulated annealing method that are believed to correspond to the global minima on the potential energy surface.

equilibrium. A cubical arrangement, for example, corresponds to a saddle-point on the potential energy surface. Consequently it is possible to deform the cube continuously to the true ground state[20], a case of spontaneous symmetry breaking. To show how the method works, I use the following simulation. Consider the restricted problem where two squares of charges lie in parallel planes a distance z above and below the equator and with relative rotation angle ϕ. The cube corresponds to $z = 1/\sqrt{3}$ $(= 0.577)$ and $\phi = 0$, and the true ground state is $z = 0.560$, $\phi = \pi/4$. Figure 1 shows the potential energy surface $E(z, \phi)$ for this constrained geometry, and the random walk towards the minimum (Fig. 1*a*). The potential energy $E(z, \phi)$ is a periodic function of ϕ, with period $\pi/2$; hence there are four equivalent minima, two of which are shown in Fig. 1. The simulation started in the right-hand side in Fig. 1*a*, at $T = 0.1$, attempted one move $(\Delta z, \Delta\phi)$ (Δz and $\Delta\phi$ are distributed uniformly in $[-0.1, 0.1]$ and $[-0.5, 0.5]$ respectively) after which the temperature was reduced by a factor of 0.9. Initially (at 'high' temperatures) almost all steps are accepted, then there is an intermediate regime where the walker is wandering on a 'ridge', and finally it walks down the 'valley' towards the minimum. Note that in a Monte Carlo simulated annealing the walker tunnels through barriers, rather than surmounting them as in a molecular dynamics simulation.

Simulated annealing is a powerful way of finding low-lying minima on a potential-energy surface, in particular for cases where steepest descent methods become too time-consuming. This has been illustrated by presenting several improvements over previous work for the three-dimensional charge-confinement problem, which is of mathematical interest[18] and may be relevant to several fields of physics[17]. Moreover, the method is very general and can easily be applied to similar problems. Results for clusters of atoms interacting under two-body Lennard-Jones forces[1] will be presented elsewhere. Even in cases where the interactions between particles cannot be expressed as a two-body force, as in quantum chemistry or solid state physics, a generalization of the method, following Car and Parrinello[25], may be possible.

Received 6 June; accepted 18 August 1986.

1. Hoare, M. R. & McInnes, J. *Adv. Phys.* **32,** 791-821 (1983).
2. Furois-Corbin, S. & Pullman, A. *Chem. Phys. Lett.* **123,** 305-310 (1986).
3. Clare, B. W. & Kepert, B. L. *Proc. R. Soc.* A**405,** 329-344 (1986).
4. Powell, M. J. D. in *Nonlinear Optimization 1981* (Academic, New York, 1982).
5. Vanderbilt, D. & Louie, S. G. *J. comp. Phys.* **56,** 259-271 (1984).
6. Wille, L. T. & Vennik, J. *J. Phys.* A**18,** L1113-L1117 (1985); *J. Phys.* A**19,** 1983 (1986).
7. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. *Science* **220,** 671-680 (1983).
8. Wille, L. T. CCP5 Inform. Q. **20,** 19-31 (Daresbury Laboratory).
9. Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E. *J. chem. Phys.* **21,** 1087-1092 (1953).
10. Berezin, A. A. *Nature* **315,** 104 (1985).
11. MacGowan, D. *Nature* **315,** 635 (1985).
12. Nityanada, R. *Nature* **316,** 301 (1985).
13. Cormack, A. M. *Nature* **316,** 301-302 (1985).
14. Queen, N. M. *Nature* **317,** 208 (1985).
15. Aspden, H. *Nature* **319,** 8 (1986).
16. Callen, M. G., Kiang, D. & Tindall, D. A. *Nature* **319,** 454 (1986).
17. Munera, H. A. *Nature* **320,** 597-600 (1986).
18. Whyte, L. L. *Am. math. Mthly* **59,** 606-611 (1952).
19. Cohn, H. *Math. Tab. natn. Res. Coun. Wash.* **10,** 117-120 (1956).
20. Goldberg, M. *Maths Comput.* **23,** 785-876 (1969).
21. Melnyk, T. W., Knop, O. & Smith, W. R. *Can. J. Chem.* **55,** 1745-1761 (1977).
22. Leech, J. *Math. Gaz.* **41,** 81-90 (1957).
23. Kellog, O. D. *Foundations of Potential Theory* (Springer, Berlin, 1967).
24. Helms, L. L. *Introduction to Potential Theory* (Wiley, New York, 1969).
25. Car, R. & Parrinello, M. *Phys. Rev. Lett.* **55,** 2471-2474 (1985).

Large AlCuLi single quasicrystals with triacontahedral solidification morphology

B. Dubost, J-M. Lang, M. Tanaka, P. Sainfort & M. Audier*

Cegedur-Pechiney, Centre de Recherche et Développement, B.P.27, 38340 Voreppe, France

* Permanent address: Laboratoire de Thermodynamique et Physico-Chimie Métallurgiques, UA CNRS 29, Domaine-Universitaire, B.P. 75, 38402 Saint Martin d'Hèref, France

The discovery of icosahedral symmetry in rapidly solidified Al-Mn alloys[1] has stimulated considerable research activity mainly because it appeared to violate the classical rules of crystallography. Hence, it left the field open for numerous reports on rapidly-solidified phases with icosahedral symmetry in aluminium based alloys. However, a challenging goal for materials scientists has always been the fabrication of large quasicrystals that would allow further investigation of the crystallographic and physical properties of this new form of material. This paper reports the first evidence of triacontahedral solidification morphology of large quasicrystalline dendrites within an as-cast Al-Cu-Li alloy with composition close to that of the Al_6CuLi_3 phase.

The first evidence of stable intermetallic phases with icosahedral symmetry, in the aluminium-lithium-copper-magnesium system[2], can be considered as an important step towards the fabrication of large quasicrystals. The intermetallic Al_6CuLi_3 compound, designated T2 (ref. 3) has been successfully cast into large dendrites surrounded by a residual eutectic, and nucleated and grown as a fine precipitate in an aluminium-rich matrix[4].

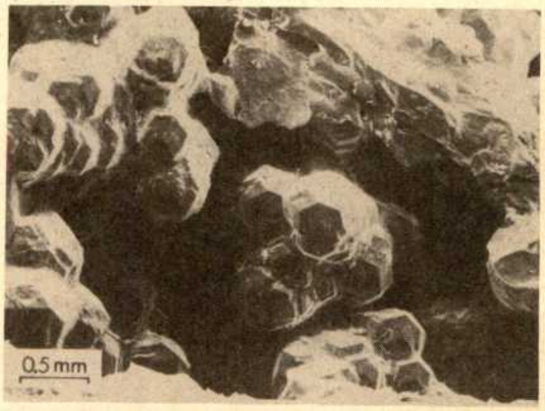


Fig. 1 Scanning electron micrograph in secondary electron mode showing the macroscopic morphology of the T2-Al_6CuLi_3 single quasicrystalline dendrites. Each dendrite consists of stacked triacontahedra with the same orientation.

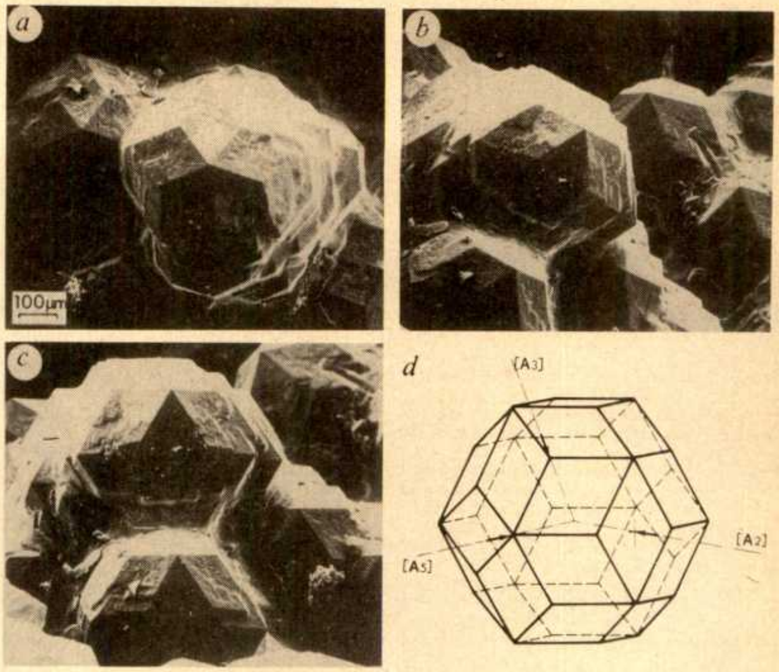


Fig. 2 Triacontahedral solidification morphology of a T2-Al_6CuLi_3 single quasicrystal viewed along a 5-fold axis (*a*), a 3-fold axis (*b*) and a 2-fold axis (*c*). These types of symmetry axes are noted on the drawing of a triacontahedron in (*d*).

These T2 phases revealed an icosahedral structure in orientation relationships with the face centred cubic matrix[4].

However, the direct production of single-phase intermetallic alloys with T2 composition proved difficult because of the occurrence of the cubic R-Al_5CuLi_3 intermetallic compound having a composition very close to that of T2-Al_6CuLi_3 (ref. 3). An inadequate alloy composition would then lead to a duplex matrix structure consisting of an R phase polyhedral core surrounded by a continuous layer of T2 (due to the non-congruent melting character of the latter phase) with orientation relationships between these 'cousin' phases[5]. We have made a major advance in this field, and detail the spectacular morphological features associated with it.

After careful weighing of very high purity aluminium, lithium and copper base metals, an alloy with chemical composition $Al_{6.4}Cu_{0.9}Li_{2.7}$ was cast at 690 °C under a controlled atmosphere in a graphite mould. The alloy was cooled slowly from 620 to 570 °C with a solidification time of 1 h, as recorded by thermal analysis. Following casting and slow cooling to room temperature, the ingot was broken open. It exhibited well defined primary dendrites along free internal surfaces, these dendrites consisting entirely of large T_2 quasicrystals to the exclusion of the other equilibrium phases.

The macroscopic solidification morphology of the T2 dendrites is of major interest: they clearly consist of stacked triacontahedra (the triacontahedron is a non-regular polyhedron with 30 rhombus-shaped faces, 32 vertices, 60 edges and icosahedral symmetry). These T2 dendrites, ~0.5 mm in diameter, can be seen in the low magnification scanning electron microscope images (Fig. 1). Their 2-, 3- and 5-fold symmetry axes, oriented perpendicularly to the picture plane, are shown in scanning electron micrographs in Fig. 2*a*, *b* and *c* respectively. They exhibit nearly perfect triacontahedral morphology (Fig. 2*d*), with the exception of a number of growth steps along some rhombic faces.

Such a triacontahedral morphology is a novel observation in crystallography and mineralogy. The triacontahedron is known to be the basic volume of the three-dimensional Penrose tiling[6,7]. This morphology can undoubtedly be attributed to the quasicrystalline nature of the primary dendrites. X-ray powder diffraction and selected area electron diffraction patterns obtained from crushed macroscopic triacontahedra (Fig. 3) reveal that the primary dendrites consist solely of T2-Al_6CuLi_3 (with no occurrence of R-Al_5CuLi_3 or other phases such as aluminium).

The striking similarity between the morphological features of these packed triacontahedra (macroscopic scale) and the atomic structure model for quasicrystals (ref. 8) recently proposed is also of great interest. This atomic structure model involves triacontahedral atomic decoration of the two rhombohedral tiles of the 3-dimensional Penrose tiling[6,7] (Fig. 4). On the other hand, the cubic Im3 structure of the R phase close to T2 (refs 9, 10) consists of a regular packing of triacontahedra[5].

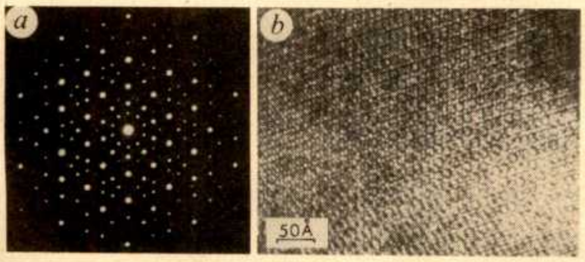


Fig. 3 *a*, Electron diffraction pattern and *b*, corresponding high resolution transmission electron microscopic image obtained along the five-fold axis from a fragment of crushed T2-Al_6CuLi_3 macroscopic triacontahedron.

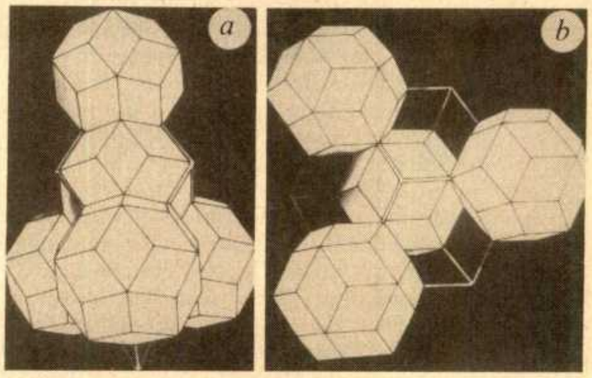


Fig. 4 Triacontahedra decorating the prolate (*a*) and oblate (*b*) rhombohedra of 3-dimensional Penrose tiling[5,8].

Using such samples, further work is now needed to describe more precisely the microscopic and geometrical features of this new morphology and to implement the structural models for quasicrystals. We propose that the results presented here on the icosahedral AlCuLi phase will lead to a better understanding and successful evaluation of the properties of this new type of material.

We thank R. Chamard, C. Bernard, P. Duroux, E. Janot and G. Duverneuil for their technical support and A. Dubus and F. X. Gilles for their contribution to previous steps of this study. Financial support of the Direction des Recherches Etudes et Techniques is gratefully acknowledged.

Received 1 August; accepted 18 September 1986.

1. Shechtman, D., Blech, I., Gratias, D. & Cahn, J. W. *Phys. Rev. Lett.* **53,** 1951–1953 (1984).
2. Sainfort, P., Dubost, B. & Dubus, A. *C.r. hebd. Séanc. Acad. Sci. Ser. II, Paris* **301,** 689–692 (1985).
3. Hardy, H. K. & Silcock, J. M. *J. Inst. Metall.* **24,** 423–428 (1955).
4. Sainfort, P. & Dubost, B. in *Workshop on Aperiodic Crystals* (eds Michel, L. & Gratias, D.) *J. de Phys. Paris, Coll. C3* **47,** 321–330 (1986).
5. Audier, M., Sainfort, P. & Dubost, B. *Phil. Mag.* (in the press).
6. Duneau, M. & Katz, A. *Phys. Rev. Lett.* **54,** 2688–2691 (1985).
7. Elser, V. *Phys. Rev.* **B32,** 4892–4898 (1985).
8. Audier, M. & Guyot, P. in *Workshop on Aperiodic Crystals* (eds Michel, L. & Gratias, D.) *J. de Phys. Paris* **47,** *Col. C3* **47,** 405–414 (1986).
9. Bergman, G., Waugh, J. L. T. & Pauling, L. *Acta crystallogr.* **10,** 254–259 (1957).
10. Cherkashin, E. E., Kripyakevich, P. I. & Oleksiv, G. I. *Sov. phys. Crystallogr.* **8,** 681–685 (1964).

Evidence for carbonate in the mantle

G. W. Berg

Geology Department, University of Port Elizabeth, PO Box 1600,
Port Elizabeth 6000, South Africa

The presence of primary carbonate in kimberlite magmas, which erupt to the Earth's surface from depths of $\geq$150 km, carrying with them the rare elemental form of carbon, diamond, is evidence for a reservoir of carbon in the Earth's mantle. Hypotheses about the form in which this carbon is stored were largely speculative until laboratory experiments[1,2] established that at depths of $\geq$50 km carbon can be incorporated in the solid carbonates, dolomite ($CaMg(CO_3)_2$) and magnesite ($MgCO_3$). However, only one instance of minute inclusions of dolomite in a sample of rock from the mantle has ever been recorded[3], despite 25 years of intensive study of these rocks. Here I report the presence of hitherto undescribed brucite/calcite intergrowths ($Mg(OH)_2/CaCO_3$) in mantle-derived xenoliths which occur in the diamond-bearing volcanic kimberlite pipes of Kimberley, South Africa. These intergrowths could reflect the presence of relatively abundant dolomite in the mantle. They seem to have been overlooked to date because of their fine textures and susceptibility to replacement by secondary serpentine.

The brucite and calcite are finely intergrown on a 10–20-μm scale (Fig. 1). Together, they form black grains of up to ~4 mm which are sometimes described in hand specimens as 'graphite' or 'carbon'. The intergrowths or their alteration products have been studied in a coarse garnet lherzolite, a porphyroclastic lherzolite and a porphyroclastic wehrlite. In the porphyroclastic samples, deformation has affected the exterior form of the black grains containing brucite and calcite, but not their microtexture (Fig. 1). The wehrlite shows about eight grains of brucite/calcite intergrowth per thin section and also contains metasomatic phlogopite. The other two samples studied contain an average of one grain of intergrowth per thin section. The collection of 70 nodules which yielded the three samples studied here is probably biased against metasomatized specimens. Thus the intergrowths may be more widely distributed in peridotite nodules from the Kimberley area than the present sampling indicates.

The brucite/calcite intergrowths are interpreted as reflecting the decarbonation products of dolomite, by analogy with the general reaction

$$CaMg(CO_3)_2 \rightarrow CaCO_3 + MgO + CO_2$$

used to describe the formation of pencatite and predazzite marbles, the periclase being hydrated to brucite[4]. The reaction in peridotite xenoliths is thought to have proceeded in response to decompression resulting from the rapid emplacement of their kimberlite hosts from depths of >100 km. The exact path of the reaction is not certain. It is known to be complicated by solid-solution effects[5] and the mantle analogue of dolomite in the equation is likely to be a solid solution between dolomite and magnesite[6]. The term dolomite is used here in this broad sense.

According to experimental data recently reviewed in ref. 7, a decrease in pressure to ~20 kbar could result in the reaction of mantle dolomite with coexisting silicates, with the accompanying evolution of CO_2 vapour. Eggler[8] has suggested that these reactions may prevent dolomite from reaching the Earth's surface. However, the observations reported here suggest that the silicate–carbonate reactions are too slow to consume the dolomite, so that the metastable silicate–carbonate assemblages persist to $\leq$5 kbar, when dolomite decarbonates[5] independently of surrounding silicates.

A wide range of oxygen fugacities has been postulated for the upper mantle[9]. The lower ranges (iron–wüstite buffer) are incompatible with the presence of carbonate. The dolomite inferred here is consistent with the EMOG (enstatite–magnesite–olivine–graphite) buffer[10].

The preservation of the texture illustrated in Fig. 1 is probably dependent on the serpentinization of the host peridotite nodule being predominantly isochemical. An abundance of circulating water would remove both the brucite formed during the serpentinization of olivine in the nodule[11], and that thought to have formed from the decarbonation of dolomite. Fairly widespread serpentinization and removal of relatively soluble brucite from assemblages of the type illustrated in Fig. 1 may explain why the occurrence of these brucite/calcite intergrowths has not been reported before.

The dolomite inferred here to be reflected in the brucite/calcite assemblages probably represents carbonate from a shallower level in the mantle than the source of the kimberlite that brought the host peridotite nodule to the surface[12]. The dolomite may be a product of mantle metasomatism, and whether such meta-

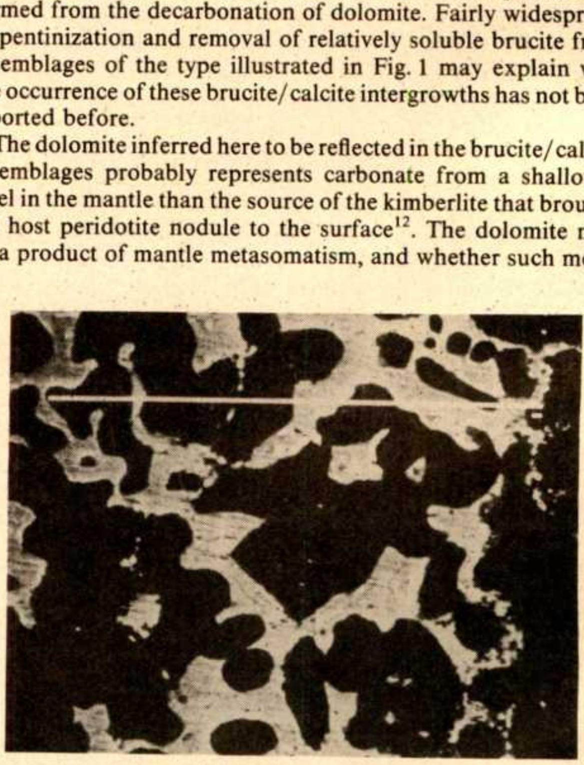


Fig. 1 Electron backscatter image of brucite/calcite intergrowth texture. Calcite (white) and brucite (black) were identified by microprobe analysis on polished thin sections and by X-ray diffraction of separated brucite/calcite grains. An X-ray raster image (not shown) shows traces of SiO_2 in the brucite which are thought to reflect incipient serpentinization. Serpentine was identified in more heavily altered brucite/calcite intergrowths by microprobe analysis. Scale bar, 100 μm.

somatism ultimately reflects plate tectonic processes or merely the interaction of magmas with wall-rock is subject to debate. However, it is suggested that the existence of the brucite/calcite assemblages described here resolves the contradiction of compelling experimental evidence for the presence of carbonate in the mantle with its apparent paucity in mantle xenoliths.

I thank R. S. Rickard and R. Moore for help with analyses on the Cameca microprobe of the Geochemistry Department at the University of Cape Town and the University of Port Elizabeth Research Committee for financial support.

Received 8 April; accepted 13 August 1986.

1. Wyllie, P. J. *J. Geol.* **86**, 687–713 (1978).
2. Eggler, D. H. *Am. J. Sci.* **278**, 305–343 (1978).
3. McGetchin, T. R. & Besançon, J. R. *Earth planet. Sci. Lett.* **18**, 408–410 (1973).
4. Harker, A. *Metamorphism* (Chapman and Hall, London, 1974).
5. Goldsmith, J. R. *J. geophys. Res.* **85**, 6949–6954 (1980).
6. Irving, A. J. & Wyllie, P. J. *Geochim. cosmochim. Acta* **39**, 35–53 (1975).
7. Wyllie, P. J., Huang, W. L., Otto, J. & Byrnes, A. P. *Tectonophysics* **100**, 359–388 (1983).
8. Eggler, D. H. *Yb. Carnegie Instn Wash.* **74**, 468–474 (1975).
9. Arculus, R. J. *A. Rev. Earth planet. Sci.* **13**, 75–95 (1985).
10. Eggler, D. H. *Geophys. Res. Lett.* **10**, 365–368 (1983).
11. Hemley, J. J., Montoya, J. W., Shaw, D. R. & Luce, R. W. *Am. J. Sci.* **277**, 353–383 (1977).
12. Wyllie, P. J. *J. geophys. Res.* **85**, 6902–6910 (1980).

Lahars initiated by the 13 November 1985 eruption of Nevado del Ruiz, Colombia

Donald R. Lowe*, Stanley N. Williams*, Henry Leigh*, Charles B. Connor†, J. Bruce Gemmell† & Richard E. Stoiber†

* Department of Geology, Louisiana State University, Baton Rouge, Louisiana 70803, USA
† Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA

The eruption of Nevado del Ruiz in Colombia on 13 November 1985 was accompanied by the formation of four lahars (mud flows) triggered by the melting of glacial ice near the summit of the volcano. The lahars began as flows of water, sand and gravel, but they incorporated clay by eroding the soil along the steep valleys through which they passed. The largest flow was a cohesive debris flow, more than 45 m deep and moving at ~12 m s^{-1} when it debouched from the canyon of the Rio Lagunilla 2.5 km from Armero, where it killed 25,000 people. The continuing volcanic activity and abundant remaining glacial ice create an extremely high risk of future destructive flows.

At 5,400 m, Nevado del Ruiz is the highest of a cluster of five volcanic peaks in the Cordillera Central of north-central Colombia (Fig. 1) and is capped by a glacier 2–5 km wide. A major eruption of Ruiz occurred in 1595, and in 1845, Ruiz was the source of a lahar which killed more than 1,000 people at Ambalema (Fig. 1) on the Magdalena River (refs 1–3 and J. B. Tomblin, unpublished work). The recent activity of Ruiz began with earthquakes in November 1984, followed by a small phreatic eruption on 22 December (refs 4, 5, 6 and J. B. Tomblin, unpublished work). A second eruption occurred on 11 September 1985 (ref. 7).

On 13 November 1985, a minor eruption at ~15:00 local time produced a light ashfall north-east of the volcano[8]. Between 21:00 and 21:15 a larger eruption occurred, producing small pyroclastic flows and surges near the summit[9,10] and melting a small part of the summit glacier, which generated four major lahars[8] (Fig. 1). The two most destructive lahars moved eastwards down the Rio Lagunilla and its tributary, the Rio Azufrado, toward Armero; a third moved northeastwards down the Rio Guali toward Mariquita; and a fourth moved northwestwards down the Rio Molinos into the Rio Claro and thence to the Rio Chinchiná toward Chichiná (Fig. 1).

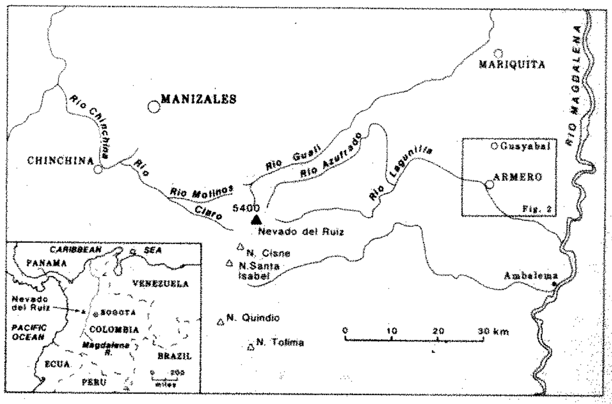


Fig. 1 Generalized location map of Nevado del Ruiz and vicinity, and (inset) Colombia.

Examination of the Rio Guali lahar deposits on the north-east slope of Ruiz, ~5 km from the summit, indicates that the initial flow was a mixture of sand, gravel, water and probably ice, but contained little clay. The main effects of the flow at this point were to strip away a 0.5–2-m-thick layer of vegetation, soil and older pumice overlying massive volcanic rock and to deposit units of coarse, moderately-to-poorly sorted, well-stratified gravel and sand. Most of the clay in the lahars that emerged from the canyons below was evidently derived by erosion of soils along the steep, V-shaped valleys through which the flows passed.

The valleys on the east and north-east side of Ruiz descend from elevations of ~5,300 m to only 400 m, where they enter the Magdalena River valley over river distances of 60–80 km. Headward slopes are generally 10–15°; slopes near the base are generally 3–4°. The lahars stripped vegetation and soil from the canyon walls to heights commonly exceeding 50 m and eroded many older alluvial deposits from the valley bottoms. The Rio Lagunilla flow travelled ~60 km before debouching just west of Armero. Based on a 2-h travel time (initiation at ~21:30 to passage through Armero at 23:30), the estimated mean flow velocity was 8.3 m s^{-1}. The course of the Rio Azufrado flow was 12 km longer than that down the Rio Lagunilla, giving a mean velocity of 10.0 m s^{-1}. Examination of the river valleys, as well as eye-witness reports, indicate that the Rio Azufrado flow was considerably more voluminous than the flow in the Rio Lagunilla, and that the former carried the bulk of the rocky debris. The differences in path-length and composition of these flows suggest that, downstream from the junction of the two rivers, the flows may have passed as a series of two or perhaps more discrete surges. Cross-cutting relationships at the junction of the Rio Lagunilla and Rio Azufrado indicate that the Azufrado flow reached the junction first[11], and may have formed the first surge to inundate Armero. The passage of a small lahar in the Azufrado in September 1985 (ref. 12) may have facilitated the movement of the later lahar.

About 1 km upstream from the canyon mouth, 2.5 km west of Armero, the Rio Lagunilla flows followed a sharp bend to the right, banking high onto the left (north) wall of the canyon (Fig. 2). The mean velocity (w) of debris flows at maximum discharge can be calculated indirectly from their superelevation around bends:

$$w = [g \cos \beta \tan \delta R]^{1/2} \qquad (1)$$

where g is the acceleration of gravity, β is the slope of the stream bed, δ is the slope of the banked flow surface (here measured from differential levels of scour on opposite sides of the river) and R is the radius of curvature of the bend[13,14].

Around this bend, $g = 9.8\ \mathrm{m\ s^{-2}}$, $\beta = 4°$, $\delta = 7.5°$ and $R = 110$ m, yielding a mean flow velocity of $11.9\ \mathrm{m\ s^{-1}}$. The flow was ~45 m deep at mid-channel and had a cross-sectional area of nearly $4{,}000\ \mathrm{m^2}$. The calculated discharge (Q) was $\sim 47{,}500\ \mathrm{m^3\ s^{-1}}$. Similar flows at Mt St Helens in the United States in 1980 had peak velocities of $3.2\text{–}40\ \mathrm{m\ s^{-1}}$ and peak discharges of 2,440 to $>190{,}000\ \mathrm{m^3\ s^{-1}}$ (ref. 15). Most lahars and debris flows attain their maximum depths near the flow fronts[13]. The conditions calculated above probably characterized the first surge of debris to leave the Rio Lagunilla canyon.

The difference between estimated mean flow velocities and that of the initial surge immediately upstream from Armero may reflect imprecise estimates of the times of flow initiation and arrival; down-slope changes in channel geometry and resistance; and down-slope flow evolution from friction-dominated gravel-sand-water slurries to cohesive gravel-sand-clay-water debris flows.

When the lahars left the canyon of the Rio Lagunilla, they emptied onto the gently sloping surface of the west side of the Magdalena River valley (Fig. 3). Within 2 km of the canyon mouth, the Rio Lagunilla is incised 10-15 m below this surface. The initial flow, however, was sufficiently deep that it spread radially from the canyon mouth and split into three main lobes. The northern lobe moved to the ENE, left the Rio Lagunilla, swept through northern Armero following a broad distributary valley, and then turned northward and flowed as a series of anastomosing lobes to within 1.5 km of Guayabal (Fig. 3).

Most of the debris flowed southeastwards, following the broad valley of the Rio Lagunilla. Because the present river lies along the southern edge of this valley, and was too small to accommodate the volume of the lahar, most of the southern lobe moved north of the river, inundating and obliterating the low-lying southern half of Armero. The only parts of the city not completely destroyed were the slightly elevated central and eastern areas between the two main flow lobes. The inertia of the flow carried all but a small part of the debris away from the present river to the south-east and east for nearly 16 km, along a broad alluvial valley which may be an abandoned river course or a route of previous lahars (Fig. 3). A much smaller amount of debris followed the course of the Lagunilla, forming a small third lahar lobe.

Early reports from Armero indicated that many victims were burned, suggesting the possibility that the lahars were hot. Although virtually everyone caught in the flow suffered extensive lacerations because of the abrasive power of the coarse debris, we saw no evidence that the flows were hot. One victim interviewed several weeks after the disaster states that the flows were cold but still burned. The high sulphate content of the flow deposits[16] suggests that the 'burning' was actually caused by the acidity of the flow.

The Lagunillas-Azufrado lahar deposits are generally between 1 and 2 m thick and mantle, rather than fill, pre-flow topography. They cover an area of nearly $32\ \mathrm{km^2}$, which, for an average depth of 1.5 m, implies a volume of $\sim 4.8 \times 10^7\ \mathrm{m^3}$. The deposits consist of grey, dense, massive, extremely poorly sorted clayey gravelly sand. Granulometric analysis of one sample from the upper part of the proximal flow deposits 0.5 km south-east of the canyon mouth shows (by weight) 47.9% sand, 27% gravel, 17.5% silt and 7.6% clay. All the coarser debris is composed of lithic volcanic rock types; fresh pumice was not identified. Clasts ≥ 50 cm in diameter are generally rounded and probably eroded from older alluvial deposits. Clasts <15 cm in diameter are commonly matrix-supported.

The grain-size structuring of the deposits indicates that, within the debris flows, the mixture of fluid and granular materials ≤ 10 cm in diameter behaved as a homogeneous phase in which there was no separation or differential settling of constituents. Clasts ≥ 15 cm across, locally reaching 10 m in diameter, occur only at the base of the lahar deposits. They were either transported at the bases of the flows or settled as the flows decelerated. As a result, many of the deposits show normal grading of the coarsest constituents.

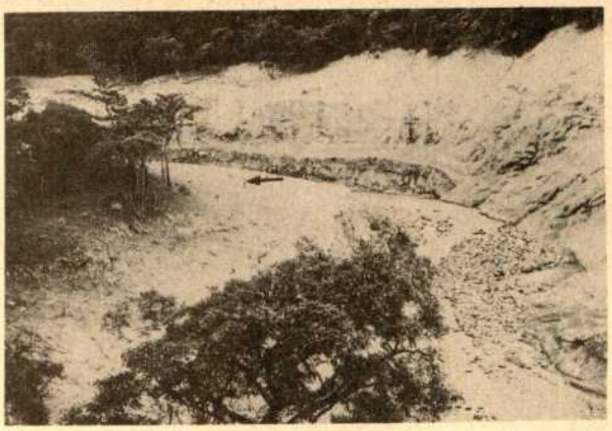


Fig. 2 Photograph of the Rio Lagunilla canyon looking upstream, ~2.5 km from Armero and 1.5 km upstream from the mountain front. Lahars generated by the 13 November 1985 eruption of Nevado del Ruiz have stripped vegetation and soil from the lower parts of the canyon walls and deposited a thin veneer of mud. A man (arrowed) gives the scale. The superelevation of the flow along the north bank (the right bank in the photograph) can be seen.

The observations that the flow deposits mantle, rather than fill, the topography and show differential settling of larger clasts suggest that the strength of the flow was relatively low. Flow strength (k) can be estimated by[13]:

$$k = T\rho_f g \sin\alpha \qquad (2)$$

where T is the deposit thickness, ρ_f is the flow density and α is the slope of the deposit surface[13]. Over broad areas, $T \simeq 100\text{–}150$ cm and $\alpha < 1.4°$, and ρ_f is taken as $2.2\ \mathrm{g\ cm^{-3}}$. For these deposits, $k \simeq (5.3\text{–}8.0) \times 10^3\ \mathrm{dyn\ cm^{-2}}$, which is comparable to that of many other debris flows[13,17].

The large volume of glacial ice remaining on Nevado del Ruiz after 13 November, the abundance of debris around the summit, and the precipitous, narrow gorges leading downward to river valleys around the volcano collectively represent ideal conditions for the formation and long-distance movement of lahars.

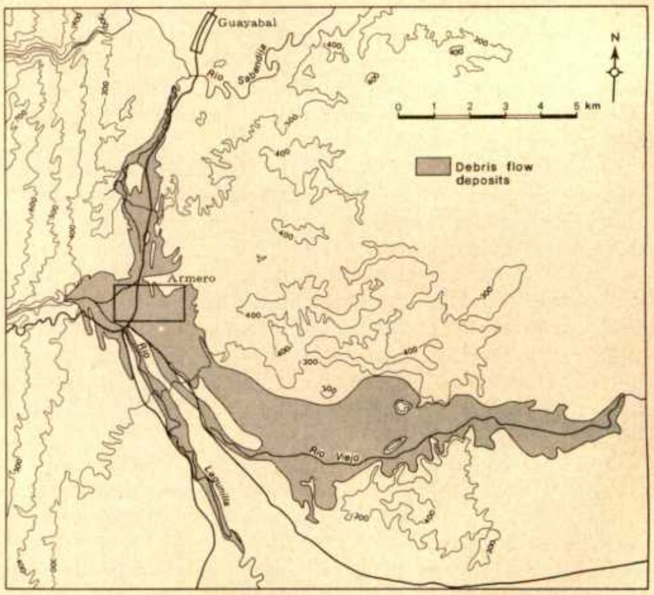


Fig. 3 Map of the Armero region (box in Fig. 1), showing the distribution of the deposits of the Rio Lagunilla lahar.

It is clear that the volcano is presently active and the probability of future eruptions is high. Even minor volcanic activity would inevitably trigger new lahars. Appropriate measures for early detection and hazard mitigation should be taken.

Supported by NSF (grant EAR86-04158 to S.N.W.), the National Geographic Society Magazine (grant to S.N.W.), the Department of Geology, Louisiana State University and the Department of Earth Sciences, Dartmouth College (Stoiber Field Fund grant to J.B.G. and C.B.C.). We thank the Fuerza Aereo Colombiana for logistical support and especially Brigadier General Melendez (FAC), NBC and ABC television, the US Army, Mr Paul Bell (US State Department), the Universidad Nacional (Manizales, Colombia), the US Geological Survey, and the following colleagues for assistance; Roberto Leigh, Omar Pedraza, Antonio Rivera, William Fritz, Hobart McDowell, Steve Raymer and Eric Johnson.

Received 14 April; accepted 13 August 1986.

1. Hantke, G. & Parodi, A. *Catalog of the Active Volcanoes of the World* Pt XIX, 2-3 (International Association of Volcanology, Rome, 1959).
2. Herd, D. G. *Glacial and Volcanic Geology of the Ruiz-Tolima Volcanic Complex, Cordillera Central, Colombia* 11 (INGEOMINAS, Bogotá, 1982).
3. Simkin, T. *et al. Volcanoes of the World* 95 (Hutchinson Ross, Stroudsburg, 1981).
4. Hall, M. L., Tomblin, J. & Goméz, O. *SEAN Bull.* **10**(5), 5 (1985).
5. Salazar-A., B., Calvache-V., M. L. & Garciá-P., N. *SEAN Bull.* **10**(7), 10 (1985).
6. Salazar-A., B. & Mejia, G. *SEAN Bull.* **10**(8), 8 (1985).
7. Jaramillo, L. *et al. SEAN Bull.* **10**(9), 3 (1985).
8. Herd, D. G. *et al. Eos* **67**, 457-460 (1986).
9. Calvache-V., M. L. *EOS* **67**, 405 (1986).
10. Janda, R. J., Banks, N. G., Pierson, T. C., Calvache-V., M. L. & Thouret, J. C. *Eos* **67**, 406 (1986).
11. Pierson, T. C., Janda, R. J., Thouret, J. C. & Calvache-V., M. L. *EOS* **67**, 406 (1986).
12. Cepeda, H. *et al. EOS* **67**, 407 (1986).
13. Johnson, A. R. *Physical Processes in Geology* 433-534 (Freeman, San Francisco, 1979).
14. Johnson, A. R. & Rodine, J. R. in *Slope Instability* (eds Brunsden, D. & Prior, D. B.) 257-362 (Wiley, Chichester, 1984).
15. Pierson, T. C. *Geol. Soc. Am. Bull.* **96**, 1056-1069 (1985).
16. Williams, S. N. *et al. Science* **233**, 964-967 (1986).
17. Johnson, A. R. & Rodine, J. R. in *Slope Instability* (eds Brunsden, D. & Prior, D. B.) 294 (Wiley, Chichester, 1984).

Sporadic shutdown of North Atlantic deep water production during the Glacial–Holocene transition?

W. H. Berger* & E. Vincent†

* Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA
† Scripps Institution of Oceanography and Département de Géologie Dynamique, Université Pierre et Marie Curie, 75005 Paris, France

Pulsations in the production of North Atlantic deep water (NADW) have been implicated in generating drastic climatic fluctuations during the Glacial–Holocene (G/H) transition[1–3]. The stable isotope record of benthic foraminifera in high-resolution cores from the Norwegian Sea suggests that such pulsations did occur[4]. Although the question of exact timing (and mechanism) is still open there is little doubt that NADW pulsations were important in climatic history because the rate of NADW production influences the rate of advection of heat to the northern North Atlantic[5]. Here we report that a sporadic shutdown of NADW may be recognizable in deep-sea carbonates with normal (low) sedimentation rates. Hence the possibility arises that relatively short-lived events (~1,000–2,000 yr) in deep circulation can be mapped over large areas of the sea floor, despite the detrimental effects of bioturbation on signal resolution.

The overall pattern of deep circulation in the present ocean is characterized by the asymmetry in deep-water properties of the Pacific and the Atlantic, which reflects the difference in age between sub-thermocline waters and, in essence, results from NADW production. We have chosen four cores to reconstruct the sequence of deep-water age difference in the two oceans since maximum glaciation (Fig. 1): ERDC 112 Bx, ERDC 123 Bx, INMD 113 Bx, and INMD 115 Bx. The ERDC box cores are from the Ontong Java Plateau in the western equatorial Pacific, from depths of 2,169 m and 2,948 m, respectively. Isotope stratigraphies and details on sedimentation rates and physical properties have been given elsewhere[6,7]. The INMD box cores are from the Mid-Atlantic Ridge, at 15.5 and 17.4° S, and from depths of 3,471 m and 3,427 m, respectively. Core data are given in ref. 8 and isotope data are listed in Table 1. Core ERDC 112 Bx and the INMD cores are close to the present two-degree potential temperature level[9]. The INMD sites are within the high-salinity-core layer of (lower) NADW. The ERDC sites sample general Pacific deep water with average salinity.

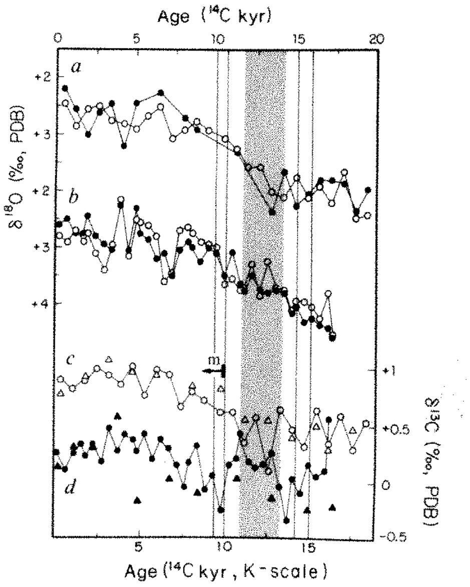


Fig. 1 Oxygen and carbon isotopes in the shells of benthic foraminifera from deep-sea sediments in the central Atlantic and western equatorial Pacific. Age scales are based on radiocarbon datings of bulk sediment (top-scale, refs 7, 8) and on correlation of planktonic oxygen isotope signals with the Norden (radiocarbon-) timescale (bottom scale, ref. 18). *a*, $\delta^{18}O$ values for *Cibicidoides wuellerstorfi* (○) and *Oridorsalis umbonatus*, adjusted by −0.6‰ (●), in core INMD 115 Bx, Atlantic. *b*, $\delta^{18}O$ values for *C. wuellerstorfi* (○) and *O. umbonatus*, adjusted by −0.75‰ (●) in core ERDC 112 Bx, Pacific. *c*, $\delta^{13}C$ values for *C. wuellerstorfi* in cores INMD 115 Bx (○) and INMD 113 Bx (△), Atlantic; m over arrow, appearance of *Globorotalia menardii*. *d*, $\delta^{13}C$ values for *C. wuellerstorfi* in cores ERDC 112 Bx (●) and ERDC 123 Bx (▲), Pacific. Atlantic data from Table 1, Pacific data from ref. 6.

The benthic foraminifera analysed are *Cibicidoides wuellerstorfi* and *Oridorsalis umbonatus*. Both species were used for plotting $\delta^{18}O$ values, but only *C. wuellerstorfi* was used for the carbon isotope record. We believe[10,11], as do others[12–16], that this species best reflects the $\delta^{13}C$ content of deep waters, with a minimum of interference from vital effects and contamination by interstitial values of $\delta^{13}C$. Analyses were carried out as described in ref. 17.

The oxygen isotope record in the deep Atlantic (INMD 115 Bx, Fig. 1*a*) shows the rise from heavy to light oxygen isotope values during the G/H transition, which is largely due to the addition of meltwater. The change occurs mostly between 14,000 and 8,000 yr BP; the G/H range is between 1.3 and 1.4‰. Of the two foraminifera, *C. wuellerstorfi* is plotted as measured, while the values for *O. umbonatus* are adjusted by −0.6‰ to make them congruent. Exact ages cannot be assigned because of the effects of resedimentation and bioturbation on radiocarbon dates; a measure of uncertainty is

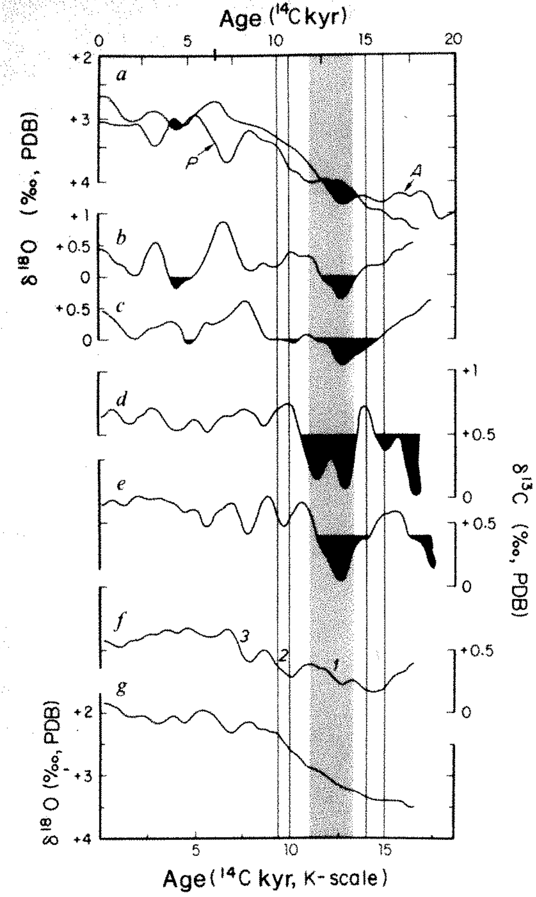


Fig. 2 Smoothed curved (triangle filter, 12321) derived from data in Fig. 1, *C. wuellerstorfi* in INMD 115 Bx (Atlantic) and in ERDC 112 Bx (Pacific). *a*, δ^{18}O values in Atlantic (A) and Pacific (P); black areas: P-values lighter than A-values. *b*, Difference in δ^{18}O values Atlantic – Pacific; black: difference negative. *c*, As *b*, after shifting Pacific curve to right by 1,000 yr. *d*, Difference in δ^{13}C values Atlantic – Pacific; black: difference $< +0.5$‰. *e*, As *d*, after shifting Pacific curve to right by 1,000 yr. *f*, Average of Atlantic and Pacific δ^{13}C curve. *g*, Average of Atlantic and Pacific δ^{18}O curve.

given by the alternative timescales shown. The timescale on top is the best-guess scale derived from bulk radiocarbon ages as listed in ref. 8. The timescale on the bottom[18] is derived from a correlation of the oxygen-isotope signal with the climate-stratigraphy of the Norden scale[19]. The stippled pattern denotes the timespan for which sporadic shutdown of NADW is proposed.

A similar oxygen isotope plot for the Pacific (ERDC 112 Bx, Fig. 1*b*) shows a noisier record (possibly because of lumpy bioturbation in this more productive area) but otherwise reflects the same transitional change. *Oridorsalis umbonatus* in this record was adjusted by −0.75‰. The greater difference between the two foraminiferal species may be a function of the higher productivity in the Pacific. The G/H range, here too, is ~1.4‰. If a range of 1.2‰ is accepted for the ice-effect[20] then the remaining 0.2‰ would suggest that glacial deep waters were roughly 1 °C colder than postglacial ones, in both oceans. The main conclusion from the oxygen isotope records is that fully glacial conditions were reached at the bottom of each of the box cores, and that the transition is well represented, apparently without hiatus.

The carbon isotope records show the Pacific/Atlantic asymmetry referred to above, with lighter δ^{13}C values in the deep Pacific, which reflects the older age of the water masses there (Fig. 1*c*,*d*). For the past several thousand years the difference in δ^{13}C is close to 0.6‰, as expected from direct measurement of water samples in the present ocean[21]. This corresponds to an age difference of ~600 yr (0.1‰ in δ^{13}C being equivalent to ~100 yr (ref. 11)). The difference persists back to ~10,000 yr BP, suggesting that there has been no major reorganization of deep circulation within this timespan. The records confirm the diminished importance of Atlantic deep-water production during glacial time, an hypothesis put forward previously on the basis of heat budget considerations[22] and benthic isotope data similar to ours[23-26].

Within a narrow time interval during deglaciation proper (the shaded zones in Figs 1 and 2) differences in δ^{13}C values sometimes disappear. We propose that this is the result of a shutdown of NADW. The decrease in Atlantic/Pacific δ^{13}C contrast during deglaciation owes as much to anomalously heavy values in the Pacific foraminifera as it does to anomalously light values in the Atlantic. It is conceivable therefore that an effect besides NADW production has to be considered in this context, namely North Pacific deep water production (see ref. 27). It is difficult to explain these anomalies as the result of bioturbation, because first, the δ-values in question are at the limit of the total range (how could mixing produce extremes?) and second, anomalous values in both Atlantic and Pacific coincide although correlation between the cores is independent of the δ^{13}C-stratigraphies.

The point at which NADW production sets in fully and permanently, at the beginning of the Holocene, is marked by the reappearance in the Atlantic of the previously regionally-extinct planktonic foraminiferan *Globorotalia menardii* (Fig. 1*c*). The coincidence between the two events may be more than fortuitous. The sinking of NADW in the Norwegian Sea implies transport of surface water across the Equator. According to Gordon[28] this water is mainly transported around the Cape of Good Hope, from the Indian Ocean, after southward flow through the Mozambique Channel. This southward flow carries tropical fauna. Thus, the beginning of NADW production favours the reseeding of the tropical foram *G. menardii* in the Atlantic, through the entrainment of tropical Indian Ocean fauna by a strengthened Agulhas Current. During glacial conditions the southward flowing current in the Indian Ocean was weakened, and colder[29,30] and was not effective in conveying *G. menardii* to the Cape. Also, entrainment around the Cape was diminished.

The δ^{13}C values show an overall Glacial to Holocene shift from light to heavy of between 0.4 and 0.5‰ (Fig. 1*c*,*d*). Such a shift is expected from the buildup of organic carbon in the terrestrial biosphere and in soils (Shackleton's 'rain forest' effect[31]). The biosphere build-up preferentially extracts isotopically-light carbon from the ocean/atmosphere system, leaving the ocean enriched in ^{13}C. A late Holocene shift towards lighter values is evident, which we ascribe to biomass destruction (desertification) following the climatic optimum.

As long as the NADW production takes place, the deep Atlantic should remain warmer than the deep Pacific[9]. If NADW production shuts off, the temperature in the deep Atlantic should fall and reach values more like those in the Pacific. To test this possibility we have superimposed the Pacific and Atlantic oxygen isotope data, after smoothing with a five-point triangle filter (Fig. 2*a*). For the past 10,000 yr the Pacific record shows the heavier (colder) values, although not the full difference expected, because of the shallower water depth of the sample cores in the Pacific (2,169 m versus 3,427 m in the Atlantic). The only time when this Atlantic/Pacific contrast reverses sign is in the early part of deglaciation, during the same interval in which the δ^{13}C difference tends to disappear sporadically. Thus, the hypothesis of NADW shutdown is supported by the oxygen isotope record.

The observed reversal of the oxygen isotope difference in the early stage of deglaciation is quite robust towards changes in timescale (that is, towards miscorrelation of the records)—see Fig. 2*b*,*c*. Figure 2*b* shows the record of the δ^{18}O difference seen

Table 1 $\delta^{18}O$ and $\delta^{13}C$ (‰, PDB) values in benthic foraminifera in INMD box cores from the South Atlantic

Depth in core (cm)	*C. wuellerstorfi*		*O. umbonatus*
	$\delta^{18}O$	$\delta^{13}C$	$\delta^{18}O$
INMD 115 Bx			
0-1	2.45	0.90	2.81
1-2	2.86	0.83	3.17
2-3	2.56	0.90	3.60
3-4	2.51	1.01	3.22
4-5	2.77	0.95	3.07
5-6	2.80	0.88	3.78
6-7	2.92	1.03	3.05
7-8	2.67	0.78	—
8-9	2.50	0.99	2.83
9-10	3.04	0.96	—
10-11	2.89	0.67	3.29
11-12	2.77	0.79	3.52
12-13	2.90	0.73	—
13-14	3.03	0.63	—
14-15	3.24	0.62	3.89
15-16	3.53	0.35	—
16-17	3.56	0.57	—
17-18	4.00	0.09	4.97
18-19	4.10	0.65	4.25
19-20	3.73	0.47	4.85
20-21	4.11	0.33	4.63
21-22	3.88	0.64	4.40
22-23	4.19	0.36	4.39
23-24	3.65	0.60	4.47
24-25	4.46	0.29	4.93
25-26	4.39	0.56	4.56
INMD 113 Bx			
0-1		0.77	
2-3		0.93	
4-5		1.08	
6-7		0.97	
8-9		0.95	
11-12		0.86	
13-14		0.82	
15-16		0.56	
17-18		0.56	
19-20		0.40	
21-22		0.50	
22-23		0.31	
24-25		0.47	

in Fig. 2*a*. If one assumes that the Pacific record should be shifted to the right by 1,000 yr to minimize the Pacific/Atlantic contrast, the reversal during early deglaciation still persists. A substantial (and unrealistic) shift of the Pacific record to the left, would be necessary to make this reversal disappear. The carbon isotope anomalies in the Pacific/Atlantic asymmetry are likewise quite robust towards miscorrelation (Fig. 2*d*,*e*).

When averaging the carbon isotope records (ERDC 112 Bx and INMD 115 Bx) an indication of stepwise change during deglaciation becomes apparent (steps labelled 1, 2, 3 on Fig. 2*f*). These steps may reflect abrupt changes in the deep-water carbon composition of the world ocean. It is interesting that they correspond to the major warming episodes proposed in the Norden scheme[19]. If these carbon shifts are due to storage of isotopically-light carbon in forests (and soils) one would then conclude that the deglacial growth of the biosphere proceeded in lockstep with the fluctuations in climate, each major warming episode producing a major increase in biosphere storage. The process would have started ~13,000 yr ago.

The average oxygen isotope record (Fig. 2*g*) shows a gradually accelerating change from the heavier glacial values to the Holocene ones, with a G/H range near 1.4‰. Substantial pre-13 kyr deglaciation[32,33] is not evident in this benthic signal. Instead, these data support a late-deglaciation model[18,34]. Such a model strengthens the case for sporadic shutdown of NADW because a high rate of meltwater influx and of warming tends to prevent the sinking and downward mixing of surface waters in the North Atlantic, during deglaciation[1].

This research was supported by NSF grant OCE85-01578. We acknowledge discussions with J.-C. Duplessy, L. Labeyrie, and N. Shackleton, on carbon isotope stratigraphy of the late Pleistocene ocean. Analyses of the $\delta^{18}O$ and $\delta^{13}C$ values were carried out at Scripps Institution of Oceanography under the supervision of J. S. Killingley.

Received 10 March; accepted 22 August 1986.

1. Olausson, E. *Progr. Oceanogr.* **3**, 221-252 (1965).
2. Berger, W. H. & Killingley, J. S. *J. mar. Res.* **40**, 27-38 (1982).
3. Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21-26 (1985).
4. Jansen, E. & Erlenkeuser, H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **49**, 189-206 (1985).
5. Berger, W. H. & Labeyrie, L. (eds) *Abrupt climatic change—evidence and implications* (Reidel, Dordrecht, in the press).
6. Vincent, E., Killingley, J. S. & Berger, W. H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **33**, 221-230 (1981).
7. Berger, W. H. & Killingley, J. S. *Mar. Geol.* **45**, 93-125 (1982).
8. Berger, W. H., Killingley, J. S., Metzler, C. V. & Vincent, E. *Quat. Res.* **23**, 258-271 (1985).
9. Mantyla, A. W. & Reid, J. L. *Deep-Sea Res.* **30**, 805-833 (1983).
10. Vincent, E., Killingley, J. S. & Berger, W. H. *Nature* **289**, 639-643 (1981).
11. Berger, W. H. & Vincent, E. *Geol. Rdsch.* **75/1**, 249-269 (1986).
12. Woodruff, F., Savin, S. M. & Douglas, R. G. *Mar. Micropaleont.* **5**, 3-11 (1980).
13. Belanger, P. E., Curry, W. B. & Matthews, R. K. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **33**, 205-220 (1981).
14. Graham, D. W., Corliss, B. H., Bender, M. L. & Keigwin, L. D. *Mar. Micropaleont.* **6**, 483-497 (1981).
15. Duplessy, J.-C. *et al. Quat. Res.* **21**, 225-234 (1984).
16. Corliss, B. *Nature* **314**, 435-438 (1985).
17. Berger, W. H. & Killingley, J. S. *Science* **197**, 563-566 (1977).
18. Berger, W. H., Killingley, J. S. & Vincent, E. *Nature* **314**, 156-158 (1985).
19. Mangerud, J., Andersen, S. T., Berglund, B. E. & Bonner, J. J. *Boreas* **3**, 109-128 (1974).
20. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39-55 (1973).
21. Kroopnick, P. M. *Deep-Sea Res.* **A32**, 57-84 (1985).
22. Weyl, P. K. *Met. Monogr.* **8**, 37-62 (1968).
23. Duplessy, J.-C., Chenouard, L. & Vila, F. *Science* **188**, 1208-1209 (1975).
24. Streeter, S. S. & Shackleton, N. J. *Science* **203**, 168-170 (1979).
25. Boyle, E. A. & Keigwin, L. D. *Science* **218**, 784-787 (1982).
26. Shackleton, N. J., Imbrie, J. & Hall, M. A. *Earth planet. Sci. Lett.* **65**, 233-244 (1983).
27. Shackleton, N. J. *Eos* **66**, 292 (1985).
28. Gordon, A. L. *Science* **227**, 1030-1033 (1985).
29. Vincent, E. *Mar. Biol.* **9**, 1-235 (1976).
30. Prell, W. L. & Hutson, W. H. *Science* **206**, 454-456 (1979).
31. Shackleton, N. in *The Fate of Fossil Fuel CO_2 in the Oceans*, 401-427 (Plenum, New York, 1977).
32. Ruddiman, W. F. & McIntyre, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 145-214 (1981).
33. Ruddiman, W. F. & McIntyre, A. *Quat. Res.* **16**, 125-134 (1981).
34. Mix, A. C. & Ruddiman, W. F. *Quat. Sci. Rev.* **4**, 59-108 (1985).

Giant subtidal stromatolites forming in normal salinity waters

Robert F. Dill*, Eugene A. Shinn†, Anthony T. Jones‡, Kevin Kelly‡ & Randolph P. Steinen§

* 610 Tarento Drive, San Diego, California 92106, USA
† Department of the Interior, Geological Survey, Fisher Island Station, Miami Beach, Florida 33139, USA
‡ University of Hawaii, Honolulu, Hawaii 96822
§ University of Connecticut, Storrs Connecticut 06268, USA

We report here the discovery of giant lithified subtidal columnar stromatolites (>2 m high) growing in 7–8 m of clear oceanic water in current-swept channels between the Exuma Islands on the eastern Bahama Bank. They grow by trapping ooid and pelletal carbonate sand and synsedimentary precipitation of carbonate cement within a field of giant megaripples. The discovery is important to geologists and biologists because similar organo-sedimentary structures built by a combination of cementation and the trapping of sediment by microbes were the dominant fossil types during the Precambrian. Stromatolites are thought to have been responsible for the production of free oxygen and thus the evolution of animal life[1,2]. Until the discovery of small lithified subtidal columnar stromatolites in the Bahamas[3], the only subtidal marine examples known to be living while undergoing lithification were

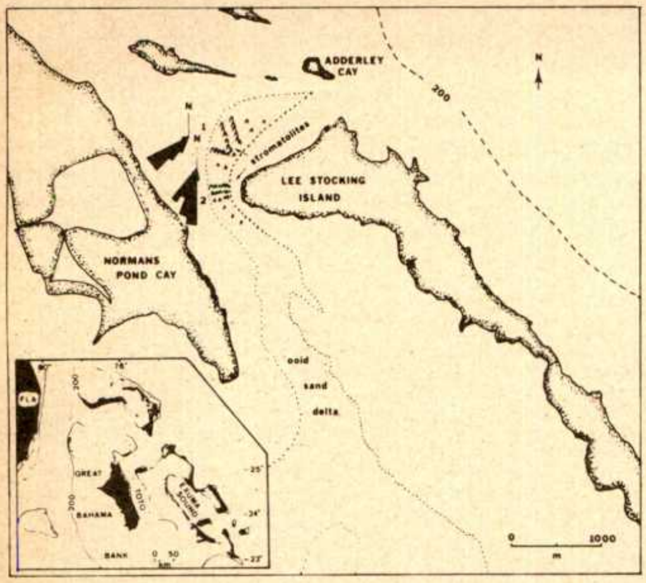


Fig. 1 Location of stromatolite-bearing tidal channel north and west of Lee Stocking Island, Exuma Cays, Bahamas. Rose diagrams are plots of the long axis of streamlined stromatolites that parallel the main axis of tidal flow in the channel: 1, 044–266° ($N=31$); 2, 014–194° ($N=60$). Orientation measurements of stromatolites were made with underwater Brunton at stations numbered 1 and 2. Visual observations demonstrate that many of the streamlined stromatolites are tilted toward the incoming tidal bottom currents. Contours in metres.

in the hypersaline waters of Hamelin Pool at Shark Bay, Western Australia[4–7]. Shark Bay stromatolites range from intertidal to the shallow subtidal with the larger columns reaching 1 m in height. The Shark Bay stromatolites have strongly influenced geological interpretation; by analogy, many ancient stromatolites have been considered to have grown in intertidal and/or hypersaline conditions[8], although hypersalinity was not a necessity for growth during the Precambrian because grazing metazoan life had not then evolved.

Stromatolites from the Exuma Islands occur in one of the many channels that exchange normal oceanic water from the deep Exuma Sound through the Pleistocene eolianite islands to the shallows of the eastern Bahama Bank. Those described here are from the channel separating Lee Stocking Island from Adderley and Normans Pond Cays (Fig. 1). The Exuma Islands are well known to carbonate geologists for their spectacular eolian carbonate dunes and ooid tidal deltas[9,10]. Exuma Sound has precipitous margins and a maximum depth of ~1.5 km. The eastern portion of Exuma Sound is open to the Atlantic Ocean and its waters which flood the stromatolites at high tide are exceptionally clear and of normal oceanic salinity (37–38‰). Underwater visibility in the tidal channel during flood tide is often >20 m. Ebb tidal waters from the shallow banks are more saline (as much as 40‰) due to evaporation, and have a visibility of 5–10 m, depending on weather conditions. Tides are diurnal and create currents between islands ranging between 60 and 100 $\mathrm{cm\,s^{-1}}$.

To date, we have found growing stromatolites only associated with migrating 1- to 2.5-m-high ooid sand dunes. The stromatolites occur as individuals or as large, coalesced bioherms that form rows perpendicular to tidal flow (Fig. 2) and as randomly-spaced individuals of variable height. Randomly-spaced individuals are usually grouped in irregular-shaped patches 5–10 m across. The tops of the larger fence-like stromatolites are parallel with the crests of submarine sand dunes. Nearly all individuals, regardless of size, are streamlined by tidal currents. Many display higher growth rates on the side facing the incoming tide, which causes them to lean noticeably towards the clear water.

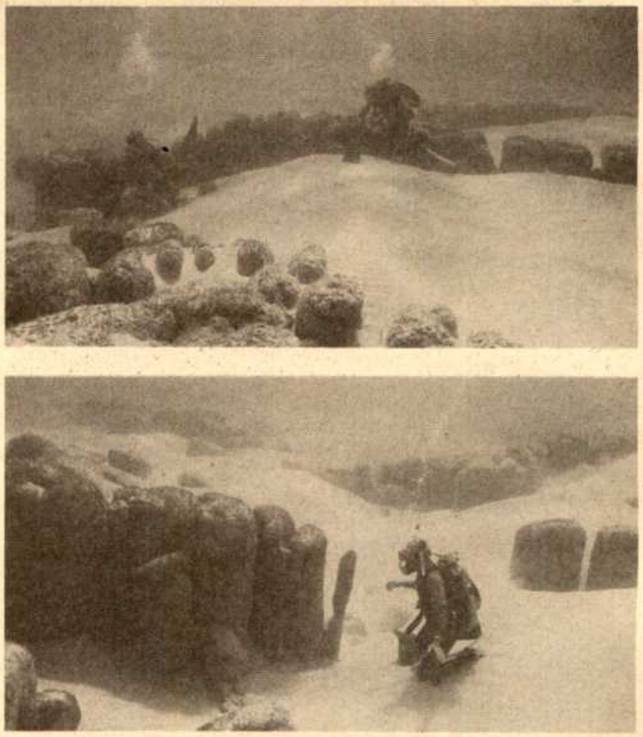


Fig. 2 Two views of large aligned fence-like stromatolites. Note that the height of migrating rippled oolitic sand dunes (megaripples) just equals the height of stromatolites and hence periodically covers most stromatolites. Smaller stromatolites in foreground of upper photograph are populated by the green alga *Batophora* spp. *Batophora* is uncommon or absent from actively growing specimens, such as in Fig. 3*a* and *c*, and those recently uncovered by migration of oolitic megaripples and bars. Photo taken at slack tide at a water depth of 7 m.

Many larger stromatolites have a soft botryoidal or pustular surface (Fig. 3*a*). Individual 1- to 10-cm-diameter botryoids are composed of uncemented ooid sand loosely bound by mucus and unidentified microbial filaments (Fig. 3*b*). Small stromatolites as much as 1 m high may also have a pustulate surface, but most are smooth and slightly elongate (streamlined) parallel to tidal currents. The surfaces of the small smooth individual are usually hard or crusty and lack the loosely held sand coating (Fig. 3*c*). Microbial filaments (Fig. 3*b*) occupy only the upper 2–4 m of the soft growth zone.

'Micro-atoll'-shaped stromatolites are also common. Micro-atolls are stromatolites with a central depression and an elevated rim, the central depressions being 10–30 cm deep and often filled with rippled oolitic sand. During maximum tidal flow, sand, shells and intraclasts swirl in the central depressions like pebbles in stream potholes. Whether the depressions are erosional features or are caused by growth of the rim has yet to be determined.

Where the channel turns southward (Fig. 1), the megaripples lose relief and the stromatolites become smaller. In this area the coalesced forms become S-shaped to form rows perpendicular to currents but with low relief, <1 m, and in plan view resemble Arabesque ornamentation. Circular micro-atolls interspersed with these shapes suggest they result from coalescence and erosion of semi-circular micro-atolls.

Staining experiments confirmed that stromatolites can grow rapidly. Three specimens were stained *in situ* with Alizarin Red S by injecting the solution into a clear inverted plastic bag cinched tightly around the base of each specimen. The bag and stain were removed after 6 h. Twenty-four h later, a layer of sand ~1 ooid thick (200 μm) had been deposited over the stained surface; sediment also adhered to vertical sides. Diver observation and underwater video documentation showed that most

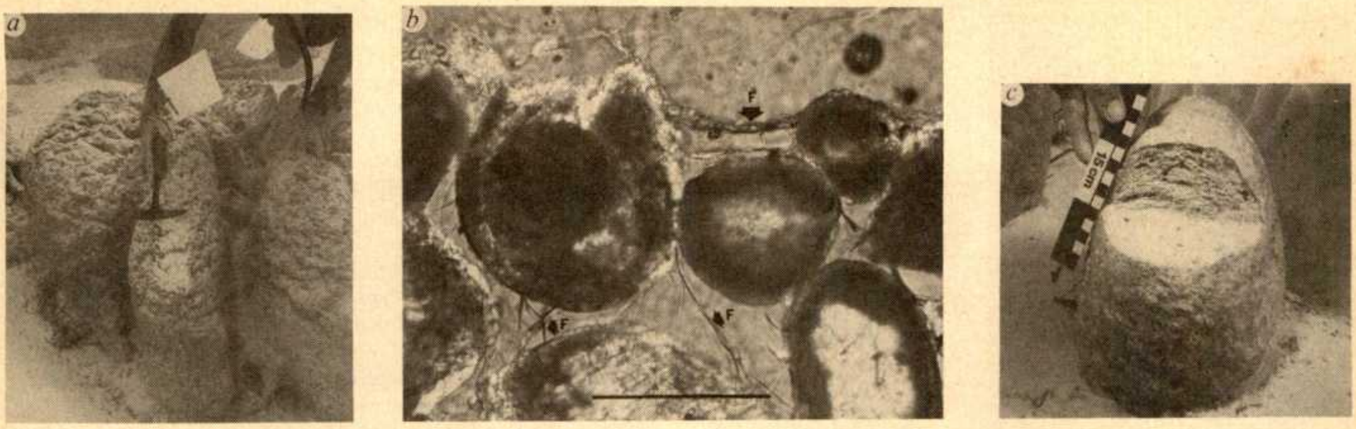


Fig. 3 *a*, Actively growing stromatolites with soft pustulate surface. *Batophora* or other epibionts are absent from surface of actively growing stomatolites. White patch below hammer caused by scraping away the soft uncemented surface layer, which is held in place by microbial filaments. Sand thrown against vertical sides by diver will stick to form a layer $\sim$1 ooid thick. *b*, Thin-section photomicrograph of growing uncemented upper surface of pustulate stromatolite. Note tiny microbial filaments (F) responsible for trapping of grains. Large grain is an ooid. Other grains are hard peloids. Note how filaments coalesce to form horizontal mat over growth surface (arrow). Growth of the internal filaments is normal to the surface. Section prepared by preserving pustule in ethyl alcohol, followed by desiccation in oven and vacuum impregnation with polyester resin. Specimen photographed under crossed polars using a gypsum plate to emphasize microbial filaments. Portion of ooid grain was plucked during preparation but algal mat shows original surface of the ooid. Scale bar, 200 μm. *c*, Underwater photograph of smooth surface of streamlined stromatolite which has been sawed with a hand saw. Note internal laminations, lack of pustules as shown in *a*, and lack of crenulations in convex-upward laminae. Ooid sand megaripple in background partially covered specimen 24 h later.

stromatolites are dusted with ooid sand during each tidal cycle and sand thrown against vertical surfaces by a diver sticks instantly to form a layer. A shell (*Strombus gigas*) from the base of a 50-cm-high specimen at a depth of 8 m gave a ^{14}C age of 480 ± 50 yr BP. The green alga *Batophora* sp., thought by Dravis[3] to be significant in stromatolite growth, was not found on the stromatolites studied here. The organic mat responsible for trapping suspended ooids and biogenic sands is a complex consortium of blue-green algae, specially adapted diatoms and as yet unidentified plants and organisms. The composition probably varies with season and time when buried columns are re-exposed. It is significant that the micro-ecosystem has adapted to high-energy conditions with life forms having 'sticky flypaper' surfaces that trap sand grains, which, in turn, on stabilization, protect the mat from further bombardment by suspended grains.

In areas near the sides of the channels, ooid sands have been stabilized by marine grasses, and sand saltation does not occur. The stromatolites in these areas are densely populated by grazing fish and epibionts such as *Acetabularia*, *Batophora*, *Sargassum*, *Siderastrea* spp. and a variety of unidentified sponges and crustose calcareous algae. Most have been invaded by boring pholad clams, and the boring sponge *Cliona* spp. has attacked many specimens; these stromatolites appear to be relict forms that are being destroyed by bioerosion.

The internal structures of dissected columns are strikingly similar to Precambrian forms and consist of convex-upward laminations composed of ooid and pelletal sand. Laminae are caused by a combination of variations in grain size, cementation and alignment of pores. Some columns have incorporated siliceous spicules. The laminations occur in two styles: nested small-scale convex-upward laminations that form 1- to 4-cm-wide columns, and larger convex-upward laminations that encompass the large portions of columnar stromatolites (Fig. 4). The smaller columns, often separated by unlaminated cemented sand, are superimposed on the larger scale laminations. Many laminations thin towards the edges of the stromatolite and in places are truncated by near vertical wall structure along the sides of the more columnar stromatolites (Fig. 4). Larger stromatolites have not been sampled, but it is thought they may be composed of both small and large columns as much as 1 m high or more. The maximum synoptic relief determined by using any one lamina in the stromatolite structure (Fig. 4) is only 16 cm, whereas the total exposed relief of the structure itself is $>$45 cm; maximum synoptic relief of the small-scale column is $<$1–2 cm.

Voids are numerous, and whereas some are borings, others are constructional. Borings, whether excavated by pholads, serpulids, or sponges, are characterized by truncated grains and cement, and except for those made by pholads, they tend to be irregular in shape and orientation. Constructional voids are planar and parallel local laminations. Grains and cement are not truncated by constructional planar voids; these can form when fleshy epiphytes such as sponges and the alga *Batophora* sp. are entombed by subsequent deposition of laminated sand. The organism decays and a void, essentially parallel to depositional laminae, is formed and preserved by rapid cementation. Finer grained sand and silt may form geopetal fillings in both kinds of voids, which often contain uncemented ooids.

Marine cementation is rapid in this ooid-forming environment, and a portion of the Bahama Bank a few kilometres west of the Exuma Islands is where the process was first documented in the Bahamas[11]. Cementation of the stromatolites is similar to that in Shark Bay[12]. Uncemented ooid sand at the surface transitionally changes to hard limestone in the interior. The gradient from organically-bound sediment to rock extends from the upper surface to a depth of 4–5 cm and is best appreciated by sawing vertically into a stromatolite with a carpenter's hand saw; sawing begins easily but slows and becomes impossible beyond 30 cm as cementation increases.

The cement is mainly acicular aragonite, similar to that observed in other areas of marine cementation. In localized areas, aragonite crystals are tightly bundled to form what has been termed acicular fan druses[13] or botryoidal aragonite[14]; the latter occurs as a partial filling in voids. Micrite-textured Mg calcite is patchy and is more common in finer grained geopetal internal sediments. Magnesian calcite also occurs as cryptocrystalline cement in 20- to 25-μm-diameter tubules[3]. Because of their habit of penetrating both pore space and grains, Dravis[15] defined the tubules as chasmolithic/interolithic algae, probably of the genus *Ostreobium*. We are unsure of their actual affinity of the timing of their entry into the stromatolite matrix. They are not common in the uncemented growth zone and thus are not considered significant as primary sediment trappers. Additional study is needed to determine when and how they invade stromatolites and become diagenetically altered to micritic Mg calcite.

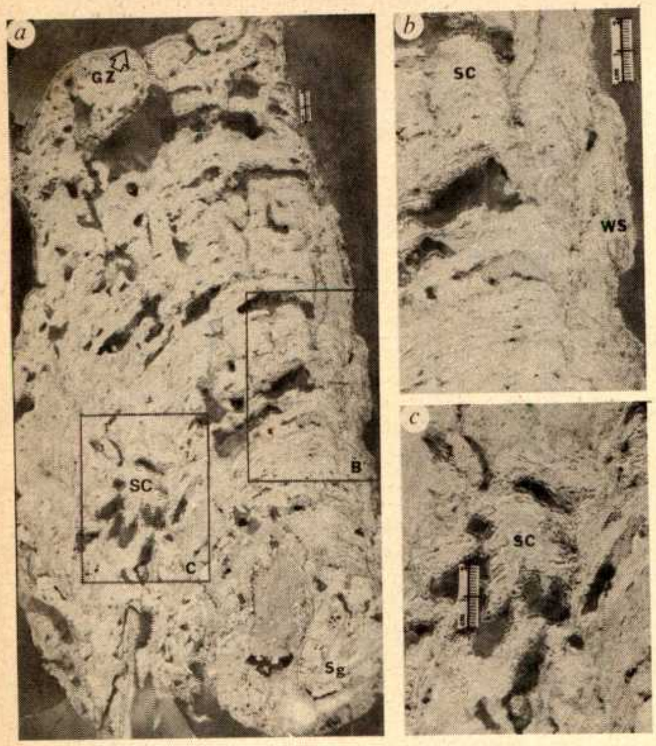


Fig. 4 Views of plastic-impregnated specimen which was cut in half to reveal internal structure. The slab is oriented so that incoming tide flowed from right to left. *a*, Slabbed plastic-impregnated specimen is 45 cm high and nucleated on a conch (*Strombus gigas*) shell (Sg). Note the small-scale convex-up laminations (SC) which form columns radiating from the base. Small-scale laminations are actually upward continuations of crenulations in larger scale laminations. Wall structure (WS) is shown in inset (*b*). Larger plastic-filled voids are borings; however, void around and below SC in inset (*c*) is thought to have been formed when a fleshy epibiont rotted away. GZ (top of *a*) shows soft uncemented pustular growth zone where grains are trapped and bound by the algal filaments shown in Fig. 3*b*. The small-scale columnar structure is thought to be caused by upward growth of pustules such as shown here. Unlaminated areas are former voids which have been filled with oolitic sand.

Bahamian subtidal stromatolites provide a new set of environmental parameters to use in interpreting spatial distribution of ancient stromatolites, orientation, water depth, salinity, constraints on growth forms and the timing of cementation and mineralization. For example, due to the Shark Bay modern analogue, most authors have regarded ancient stromatolites as intertidal in origin, although some have subtidal affinities or large size, which necessitated postulating enormous paleotidal fluctuations[8]. Modern subtidal stromatolites have been reported before[3,6,16,17] but unfortunately seem to have had little influence on the mainstream of geological interpretation. The Bahamian examples described here are by far the deepest and largest to be reported from the Holocene and hopefully may help clarify the many debates concerning salinity, current regimen, and depositional water depth of ancient stromatolites. Ancient stromatolites, like modern ones, were not limited to intertidal or hypersaline conditions, and many authors recognized this as a fact based on geological interpretation[8,18–20]. The Bahamain stromatolites provide an alternative for palaeoenvironmental comparative studies.

In the Holocene and probably throughout much of Phaerozoic time, grazers and borers prevented stromatolite growth and preservation in normal marine subtidal waters. During the Precambrian, however, there were no known grazing and boring organisms, and stromatolites flourished uncontested for $>2\times10^9$ yr. Without competition, Precambrian stromatolites freely populated enormous areas[20,21], and probably grew in water depths >10 m (ref. 8). Strong tidal currents were not a necessary requirement for growth and preservation in the Precambrian before the evolution of higher life forms. We anticipate that future comparative research on these Holocene stromatolites will provide many additional clues for interpreting depositional environments and the processes that moulded the shapes and distribution of ancient stromatolites of all ages.

We are grateful for the support and facilities at Lee Stocking Island provided by the Caribbean Marine Research Center. We acknowledge the support of Mr Robert Wicklund, Ms Geraldine Wenz, and Mr John H. Perry Jr, owner of Lee Stocking Island. Partial support was provided by the Office of Undersea Research, National Oceanic and Atmospheric Administration, US Department of Commerce. Ralph Knight is thanked for piloting us. We thank Richard Wilkens, Thomas Lorensen and Tom Judy for underwater support, field and laboratory help in this unfunded research. Barbara Lidz provided editorial assistance. We also thank Sal Valestro at the Balcones Research Center of the University of Texas for dating the conch shell and R. N. Ginsburg, Richard Slater and Christopher Kendall for field assistance and stimulating discussion. Finally, we thank J. H. Hudson for guidance, equipment and chemicals for staining stromatolites *in situ* and Renato Diaz for sample and thin section preparation, and photoprocessing.

Received 11 June; accepted 22 August 1986.

1. Rutten M. G. *The Geological Aspects of the Origin of Life on Earth* ((Elsevier, Brussels, 1962).
2. Berkner, L. V. & Marshall, L. C. in *The Origin and Evolution of Atmospheres and Oceans* (eds Brancazio, P. J. & Cameron, A. G. W.) 102–126 (Wiley, New York, 1964).
3. Dravis, J. J. *Science* **219,** 385–386 (1983).
4. Logan, B. W. *J. Geol.* **69,** 517–533 (1961).
5. Logan, B. W., Rezak, R. & Ginsburg, R. N. *J. Geol.* **72,** 68–83 (1964).
6. Playford, P. E. *Western Aust. Geol. Survey A. Rep. 1979* 73–77 (1980).
7. Awramik, S. M. & Vanyo, J. P. *Science* **231,** 1279–1281 (1986).
8. Monty, C. in *Fossil Algae, Recent Results and Developments* (ed. Flügel, E.) 15–35 (Springer, Berlin, 1977).
9. Illing, L. V. *Am. Ass. Petrol. Geol. Bull.* **38,** 1–95 (1954).
10. Ball, M. M. *J. Sedim. Petrol.* **37,** 556–591 (1967).
11. Taft, W. H., Arrington, F., Haimoritz, A., MacDonald, C. & Woolheater, C. *Bull. Mar. Sci. Gulf Caribbean* **18,** 762–828 (1968).
12. Burne, R. V. & James, N. P. *12th IAS Congr.* Canberra (in the press).
13. Shinn, E. A. *Sedimentology* **12,** 109–144 (1969).
14. Ginsburg, R. N. & James, N. P. *Geology* **4,** 431–436 (1976).
15. Dravis, J. J. *J. Sedim. Petrol.* **49,** 0195–0208 (1979).
16. Ginsburg, R. N. *Proc. int. Geol. Cong. XXI, Norden 1960,* pt XXII (Copenhagen) 26–35 (1960).
17. Gebelein, C. D. *J. Sedim. Petrol.* **39,** 49–69 (1969).
18. Playford, P. E. & Cockbain, A. E. *Science* **165,** 1008–1010 (1969).
19. Hoffman, P. E. *Am. Ass. Petrol. Geol. Bull.* **58,** 856–867 (1974).
20. Cloud, P. E. *Am. J. Sci.* **240,** 363–379 (1942).
21. Walter, M. R. *Spec. Pap. Palaeontol.* No. 11 1–190 (Palaeontological Association of London, 1972).

Sex-ratio manipulation in the common opossum

Steven N. Austad & Mel E. Sunquist*

Department of Organismic and Evolutionary Biology,
The Biological Laboratories, Harvard University, Cambridge,
Massachusetts 02138, USA
* Florida State Museum, University of Florida, Gainesville,
Florida 32611, USA

The idea that offspring sex ratio is affected by parental capacity for reproductive investment was first developed by Trivers and Willard[1] and later elaborated by others[2–6]. According to this hypothesis, if the amount of parental investment[7] differentially affects the expected reproductive success of offspring depending on offspring gender, then parents capable of exceptionally large investment should bias reproductive investment toward the sex affected most. For most polygynous mammals, male reproductive success is expected to be most enhanced by exceptionally large investment[1,8], but observational field studies and laboratory experiments have either supported[9–13] or contradicted[3,14–16] this conditional sex-allocation hypothesis. We have performed the first

experimental field study specifically to examine this hypothesis, and find that the female common opossum (*Didelphis marsupialis*), whose capacity for parental investment was artificially enhanced by dietary supplementation, produces male-biased sex ratios, in contrast to the unbiased ratios produced by control females.

The conditional sex-allocation hypothesis holds that variance in reproductive success is expected to be greater among males than among females in polygynous animals because some males have more than one mate whereas other males have no mates. Consequently, because high-quality males will have greater reproductive success than any single female, females producing high-quality offspring will have more grand offspring if they produce males. Conversely, low-quality offspring should be female, because essentially all females find mates.

The common opossum, a polygynous marsupial[17], exhibits no paternal care. Young are born after 13 days' gestation in an embryonic state and complete development during 60–70 days in the maternal pouch. During a 2-year study near Calabozo in the central Venezuelan savanna, we trapped, weighed, radio-collared and released 40 adult virgin females (virgins can be identified by the condition of their pouch[18]) before the onset of the breeding season. Half of these females were randomly assigned to control and half to experimental groups. Two controls died before the start of breeding and are not included in the results. Experimental animals were located in their day time sleeping dens and provisioned with ~125 g sardines every 2 days from at least 2 weeks before any pouch young appeared until the end of the breeding season. Provisions were placed at the mouth of their dens near dusk when the animals become active. Provisioning was intended as a supplement to, and not a substitute for, a normal diet. Both controls and provisioned animals were recaptured, weighed and the size and sex of their pouch young recorded approximately once per month.

Pouch young can be toe-clipped and sexed visually at about 10–15 days of age when they weigh only ~1 g (<0.1% of maternal body weight). The course of pouch-young development in opossums has been extensively studied[18–20], so that from a combination of anatomical features the animals can be aged to within 5 days for their first month in the pouch.

We found that provisioned females have no more offspring than controls (Table 1), but their individual offspring are significantly larger at all stages of development from 20 days of age until they left the maternal pouch (Fig. 1). Thus, provisioned females exhibit greater total parental investment than controls. The recapture rate of independent juveniles supports the contention that the offspring of provisioned females survived better than controls (Table 1).

Table 1 Maternal investment in control versus provisioned females

	Control	Provisioned	P-value
Maternal weight (kg) at start of experiment ($\bar{X}\pm$s.e.m.)	1.38±0.07	1.37±0.07	0.71 ($t_{36}=0.314$)
Litter size ($\bar{X}\pm$s.e.m.)	7.52±0.36	7.61±0.34	0.56 ($t_{36}=0.142$, one-tail)
Pouch young mortality (%) from age <15 days to >45 days			
Males	2.3 (87)	4.5 (111)	0.30
Females	2.3 (87)	5.2 (77)	0.22
Total	2.3 (174)	4.8 (188)	(G test)
Weaned young recaptured (%)			
Males	16.4 (55)	44.4 (54)	0.001
Females	12.8 (47)	28.9 (45)	0.054
Total	14.7 (102)	37.4 (99)	0.0002 (G test)

Numbers in parentheses, total number in sample.

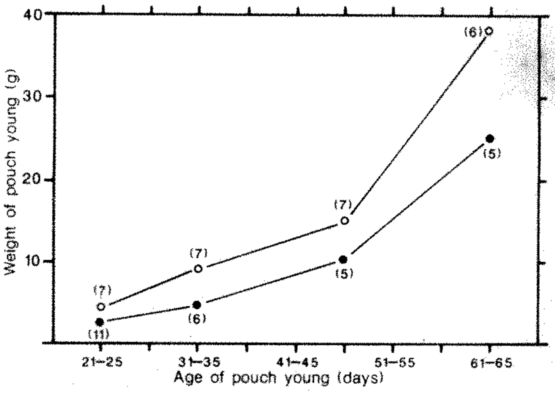


Fig. 1 Difference in weight of pouch young between provisioned (○) and control (●) females. Data from all 5-day periods during which at least 5 females from both groups were captured. Weight differences are statistically significant for all comparisons (Wilcoxon rank sum tests: 21–25 days, $P=0.038$; 31–35 days, $P=0.007$; 46–50 days, $P=0.015$; 61–65 days, $P=0.026$). No. of females in each sample in parentheses. Fifteen females from both provisioned and control groups are represented. Samples at 46–50 and 61–65 days are independent in that no female is represented in both samples.

Reproductive allocation can be sexually biased even with an unbiased sex ratio if investment is greater in individuals of one sex[21–23]. Although there is often marked variation in size among litter mates, we found no evidence of skewed somatic investment in individuals of either sex. Mean weights of male and female litter mates are statistically indistinguishable (Wilcoxon signed-rank test for paired observations, $P>0.2$ for controls and $P>0.5$ for provisioned females). Neither is there a sexual bias in mortality of pouch young examined between 15 and 45 days of age (Table 1) in either provisioned or control groups (G test, $P>0.8$ for both groups). Although maternal investment per individual is unbiased in both groups, provisioned females invest disproportionately in males by producing male-biased sex ratios. Controls, on the other hand, produce exactly as many male as female offspring (Fig. 2). The sex-ratio difference between groups is statistically significant: P (one-tail) = 0.007, binomial test. Overall, male-biased litters were produced by 55% of control females and 76% (13 of 17) of provisioned females with a sex-ratio bias.

The mechanism of this apparent sex-ratio manipulation is unknown. The sex ratios described above are those observed when pouch young were first old enough to be visually sexed. However, <2% of pouch young attached to teats died before the age at which they could be sexed, indicating that the sex ratios observed were essentially the same as those of initially attached young. Female opossums typically bear more young than can make their way to the pouch and find a functional teat[18], so it is possible that males of provisioned mothers are superior to females in their capacity to migrate successfully to the pouch. Alternatively, adult females might distinguish male from female young at birth and selectively remove females as they travel toward the pouch. Another possibility is that sex-ratio manipulation occurs before birth.

For this particular sex-ratio manipulation to be an adaptive response, the amount of maternal investment would have to have a larger effect on male than female reproductive success. One reason to suspect that this may be the case is the tendency for improvement in recapture rate of independent young from provisioned versus control females to be greater for males than females ($P=0.15$, binomial contrast; Table 1), suggesting that male survival is more affected by increased investment than is female survival. Other support for a greater effect on future male reproductive success derives from the putative relationship between adult body size and reproductive success. Among our control females, body size is uncorrelated with such fitness components as length of adult life ($r=0.088$, $t_9=0.250$, $P>0.5$),

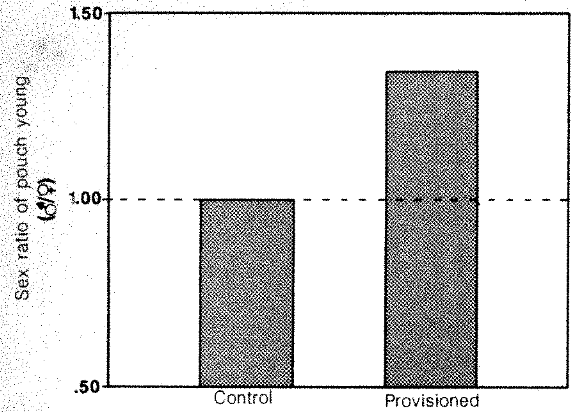


Fig. 2 Sex ratios of pouch young from control and provisioned females. Sex-ratio difference is significant (P(one-tail) = 0.007, binomial test). Control females, $n = 256$; provisioned females, $n = 270$.

litter size ($r = 0.135$, $t_{16} = 0.670$, $P > 0.5$), number of litters produced ($r = -0.065$, $t_9 = 0.195$, $P > 0.5$) or size of individual offspring ($r = -0.370$, $t_9 - 1.19$, $P > 0.25$). On the other hand, male body size could well be related to mating success. Male common opossums have large overlapping home ranges, so that each male has potential access to multiple females, but must contend with many other males who also have potential access[17]. Males fight vigorously in the presence of a receptive female, and a high proportion of all males captured during the breeding season show wounds consistent with male-male combat (our unpublished observations), whereas females do not exhibit such wounds. Male body size is known to be correlated with the ability to win contests in some mammals[24-26] including marsupials[27,28], and therefore plausibly correlates with access to mates.

Documentation of adaptive sex-ratio manipulation among mammals has remained elusive[2,6,8,29]. Part of the difficulty in observational studies could be in identifying key variables correlated with capacity for parental investment. For instance, the offspring sex ratio in red deer varies adaptively with maternal dominance rank[12], but does not vary with other plausible variables such as maternal age, home range size or parity[30]. Environmental variation could also obscure significant trends during any field study; for example, in a similar study of the Virginia opossum, food supplementation resulted in an overproduction of males during a year of normal rainfall, but failed to do so in a drought year (M.E.S., in preparation), possibly because even provisional females were not in exceptionally good condition during the drought. However it seems increasingly clear that taxonomically diverse mammal species can adjust their sex ratios as the theory predicts.

We thank Tomás Blohm for allowing us to work on his ranch and assisting us with logistical problems. We also thank Dennis Daneke and his assistants Brent Mitchell and David Manry for performing most of the fieldwork. We thank E. L. Charnov, T. H. Clutton-Brock, G. D. Dodson, J. F. Eisenberg, L. D. Marshall, J. Maynard Smith, W. R. Rice, M. J. A. Simpson and R. Thornhill for helpful comments on the manuscript. Supported by NSF grant BSR-8315125.

Received 28 March; accepted 21 August 1986.

1. Trivers, R. L. & Willard, D. E. *Science* **179,** 90–92 (1973).
2. Williams, G. C. *Proc. R. Soc.* **B205,** 567–580 (1979).
3. Altmann, J. *Baboon Mothers and Infants* (Harvard University Press, 1980).
4. Fiala, K. L. *Evolution* **35,** 898–911 (1981).
5. Okansen, L. *Am. Nat.* **117,** 109–111 (1981).
6. Charnov, E. L. *The Theory of Sex Allocation* (Princeton University Press, 1982).
7. Trivers, R. L. in *Sexual Selection and the Descent of Man 1871–1971* (ed. Campbell, B.) 136–179 (Aldine, Chicago, 1972).
8. Clutton-Brock, T. H. & Albon, S. D. in *Current Problems in Sociobiology* (ed. King's College Sociobiology Group) 223–247 (Cambridge University Press, 1982).
9. Rivers, J. P. W. & Crawford, M. A. *Nature* **252,** 297–298 (1974).
10. McClure, P. A. *Science* **211,** 1058–1060 (1981).
11. Gosling, L. M. & Petrie, M. in *Physiological Ecology* (eds Calow, P. & Townsend, C. R.) 315–345 (Blackwell Scientific, Oxford, 1982).
12. Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. *Nature* **308,** 358–360 (1984).
13. Rutberg, A. T. *Am. Nat.* **127,** 89–94 (1986).
14. Verme, L. J. *J. Wildl. Mngt.* **47,** 573–582 (1983).
15. Simpson, M. J. & Simpson, A. E. *Nature* **300,** 440–441 (1982).
16. Silk, J. B. *Am. Nat.* **121,** 56–66 (1983).
17. Sunquist, M. E., Austad, S. N. & Sunquist, F. *J. Mammal.* (in the press).
18. Petrides, G. A. *J. Mammal.* **30,** 364–378 (1949).
19. Reynolds, H. C. *Univ. Calif. Publs Zool.* **52,** 223–284 (1952).
20. Gardner, A. L. in *Wild Mammals of North America* (eds Chapman, J. A. & Feldhamer, G. A.) 3–36 (Johns Hopkins University Press, 1982).
21. Reiter, J., Stinson, N. L. & LeBoeuf, B. J. *Behav. Ecol. Sociobiol.* **3,** 337–367 (1978).
22. Dittus, W. P. J. *Behaviour* **69,** 265–301 (1979).
23. Maynard Smith, J. *Behav. Ecol. Sociobiol.* **7,** 247–257 (1980).
24. Geist, V. *Mountain Sheep* (University of Chicago Press, 1971).
25. Clutton-Brock, T. H., Guinness, F. E. & Albon, S. D. *Red Deer: Behaviour and Ecology of Two Sexes* (Cambridge University Press, 1982).
26. Fleming, T. H. *J. Mammal.* **55,** 543–561 (1974).
27. Kaufmann, J. H. *Am. Zool.* **14,** 51–62 (1974).
28. Johnson, C. N. *Austral. Wildl. Res.* **10,** 25–31 (1983).
29. Myers, J. H. *Am. Nat.* **112,** 381–388 (1978).
30. Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. *Nature* **289,** 487–489 (1981).

Wormy mice in a hybrid zone

Richard D. Sage*†, Donald Heyneman‡,
Kee-Chong Lim‡ & Allan C. Wilson*

* Department of Biochemistry and † Museum of Vertebrate Zoology,
University of California, Berkeley, California 94720, USA
‡ Department of Epidemiology and International Health,
University of California, San Francisco, California 94143, USA

As one approach to analysing the genetic barriers between species, we studied the numbers and types of parasitic worms in two species of house mice (*Mus musculus* and *M. domesticus*) and in their natural hybrids. Where the ranges of these two species meet in southern Germany, there is a zone of hybridization less than 20 kilometres across[1], in which about 98% of the mice have backcross genotypes. Fourteen of the 46 mice tested from within the zone have over 500 pinworms per gut, a number far exceeding the mean of 40 per gut for other mice inside and outside the zone. Other nematodes have a similar, non-random distribution. The number of mice bearing 9 or more tapeworms per gut is also excessive in the hybrid zone. These extraordinarily wormy mice may be unusually susceptible to parasitism; the different species may have different genes for resistance, and recombinant backcross animals may lose both[2]. Our findings support the view that the hybrid populations may have reduced fitness and thereby act as a genetic sink, interfering with the flow of genes between the two species[1,3]. The possibility that environmental or ecological peculiarities in the zone of hybridization make the mice more liable to infection is not supported.

Ninety-three mice, caught in barns and houses at 22 localities in southern Germany and adjacent Austria, were necropsied for worms in their tissues and intestinal tracts. These mice are part of a larger collection of about 300 individuals from 30 localities along a 200-km transect extending between Braunau and Augsburg. This transect encompasses the reported zone of hybridization between the eastern *M. musculus* and western *M. domesticus*[4,5]. A partial genetic characterization of mice from these 30 localities has appeared[1]. Animals were trapped alive during October 1984 (42 mice) and June 1985 (51 mice) and killed within 2 days of capture. All but 5 of the 93 mice were adults, weighing more than 12 g, and 40 were males. The carcasses were preserved in 10% neutral-buffered formalin for 1–2 months, soaked in water, and stored in 70% ethanol. The entire gastrointestinal tract was slit and examined, with associated organs, under a dissecting microscope.

The genotype of each mouse at four enzyme loci was determined by starch gel electrophoresis[1]. These two species have

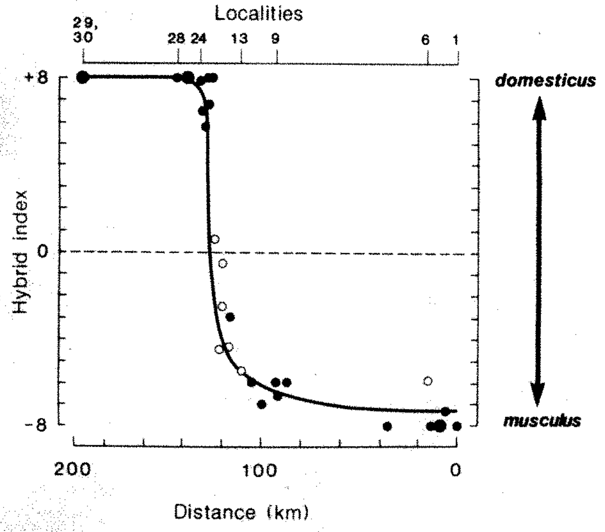


Fig. 1 Mean hybrid indices of 30 mouse populations based on electrophoresis of four diagnostic proteins. The samples come from localities 1–30 lying along a line extending between Braunau, Austria (1), and Augsburg, Germany (29, 30; ref. 1). Locality numbers of some of the 22 places from which mice were autopsied for parasites are indicated across the top; larger circles, two collections from the same city; ○, localities at which the 15 mice with excessive burdens of nematodes were found.

fixed allelic differences at loci for esterases 1 and 10, mannose phosphate isomerase-1 and nucleoside phosphorylase-1. A hybrid index (HI) based on the genotypes at these four loci was computed for 200 mice, including the 93 animals autopsied for parasites. As shown in Fig. 1, there is a nearly complete replacement of one species by the other across a distance of less than 20 km. The mitochondrial DNA, chromosomes, osteology, pelage coloration and hybrid sterility alleles also show this abrupt transition (R.D.S., unpublished).

Table 1 Numbers of worms per gut in 93 wild mice

Locality*	Hybrid index	No. mice	Mean abundance (maximum) Nematodes†	Cestodes‡
		Eastern region (*M. musculus*)		
1	−8	3	4.3 (10)	1.7 (3)
2	−7.3	2	7.5 (15)	0
4	−8	2	16 (30)	0
6	−6.0	8	100+ (500+)	1.4 (3)
9	−6.6	4	80 (151)	0.8 (1)
12	−6.0	2	25 (30)	0
		Hybrid zone		
13	−5.5	2	255+ (500+)	0
14	−3.0	2	20 (40)	0
15	−4.3	5	423+ (520+)	11 (22)
16	−0.5	4	170+ (500+)	0.6 (2)
17	−4.5	7	79+ (500+)	0
18	−2.5	3	350+ (500+)	3.0 (9)
19	+0.6	13	243+ (620+)	5.8 (30)
20	+6.7	4	44 (100)	1.0 (4)
23	+5.7	6	126 (200)	0
		Western region (*M. domesticus*)		
24	+6.5	7	2.0 (6)	0
25	+7.8	4	22 (30)	0
26	+8	3	8.3 (14)	0
27	+8	1	8	11
28	+8	2	28 (52)	0.5 (1)
29	+8	6	8.3 (15)	0.5 (2)
30	+8	3	54 (100)	0

* The locality numbers and the mean hybrid indices are given in ref. 1, Table 1.

† The five genera represented and total numbers of specimens found are: *Aspiculuris* (>5,740), *Syphacia* (>3,865), *Protospirura* (485), *Trichocephalus* (512) and *Trichuris* (65).

‡ The principal cestode genera (and the total numbers found) were *Hymenolepis* (97), *Catenotaenia* (64) and *Taenia* (14).

We found and identified more than 10,840 worms in the 93 mice. From Table 1, which gives the average number of worms per mouse at the 22 localities, it is evident that the highest worm burdens occur in the hybrid zone. Here, between localities 13 and 23, the average number of nematodes per mouse is greater than 197, the standard error being 32. By contrast, in the east the mean is 59 (±23) and in the west it is only 15 (±4). The caecal pinworms (*Aspiculuris* and *Syphacia*), which comprised 90% of the nematodes (Table 1), and other genera (*Protospirura*, *Trichocephalus* and *Trichuris*) also occurred in highest numbers in hybrid mice. A similar trend is apparent for cestode infections (Table 1): mice in the hybrid zone have an average of 3.2 (±0.97) worms, whereas the eastern and western mice have 0.9 (±0.23) and 0.6 (±0.13) respectively.

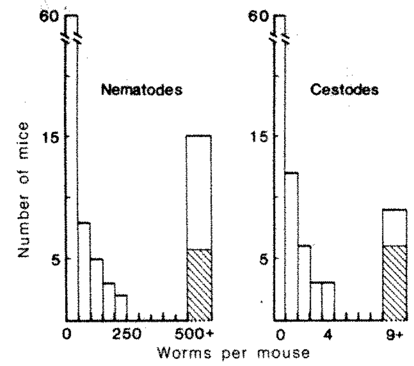


Fig. 2 The distribution of parasite burdens (number of worms per mouse) among 93 wild house mice captured at the 22 localities cited in Fig. 1: *a*, nematodes *b*, cestodes. Hatched columns, six mice that had excessive numbers of both nematodes (>500) and cestodes (>9).

The number of worms per mouse varies widely, from none to >620 (Table 1, Fig. 2). Table 2 shows that the distributions of worm burdens per mouse are highly skewed for each of the three regions, the variance exceeding the mean in nearly every case. Such skewed distributions are characteristic for parasites within their host populations[6,7]. The ratio of the variance to the mean is highest for the hybrid zone. This observation, in conjunction with the evidence shown in Fig. 3 and discussed below, may imply an association between worm numbers and particular genotypes.

We designate the 15 mice bearing >500 nematodes as 'wormy' and the remaining 78 mice (<250 nematodes) as 'normal' (most of the normal mice carry <50). The pattern of nematode burdens is associated with particular protein genotypes (Fig. 3). Fourteen of the 15 wormy mice have hybrid genotypes. Only one of these heavily parasitized mice has the genotype of a pure *M. musculus.* A χ^2 analysis of pure versus hybrid genotypes and wormy (>500 nematodes) versus normal (<250) mice gives a highly significant result ($\chi^2 = 9.6$, d.f. = 1, $P < 0.005$). Thus there are significantly fewer pure mice, and more hybrids than expected, that are heavily parasitized with nematodes.

Cestode infections show a similar pattern (Fig. 2*b*), although the maximum number of parasites per mouse is not as great as for nematodes. There are 84 normal animals with from zero to four cestodes, and nine wormy mice with heavy infections ranging from 9 to 30 worms. Eight of these wormy mice have hybrid genotypes (HI = −4 to +4), while one pure *domesticus* mouse had 11 cestodes. A χ^2 test of pure versus hybrid genotypes and

Table 2 Statistical analysis of worm numbers per mouse

Type of worm	Variance divided by the mean Eastern region ($n = 21$)	 Hybrid zone ($n = 46$)	 Western region ($n = 26$)
Nematodes			
Syphacia	263	368	50
Aspiculuris	51	343	30
Protospirura	3.9	84	1.0
Trichurids	9.1	160	18
Cestodes			
Hymenolepis	6.9	25	10
Other taenioids	1.5	12	0.8

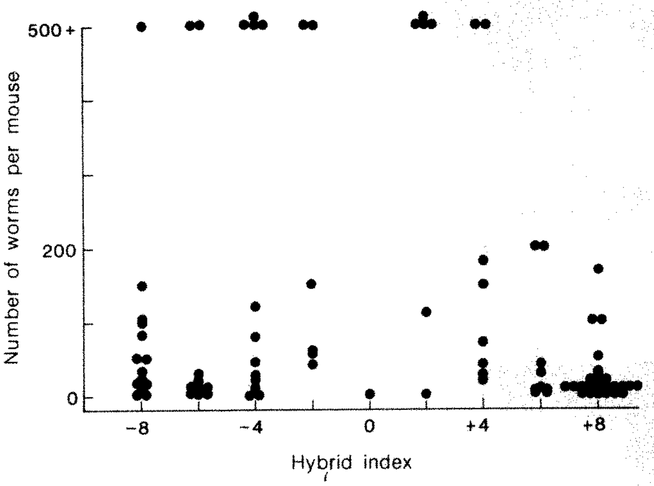


Fig. 3 Nematode burdens (number of worms per mouse) in relation to mouse genotype. The hybrid index is based on electrophoresis of four diagnostic proteins. Pure *M. musculus* has an index of −8; pure *M. domesticus* has an index of +8.

wormy (>8 cestodes) versus normal (<5) mice gives a significant result ($\chi^2 = 4.1$, d.f. $= 1$, $P < 0.05$). We infer that hybrid mice can be more heavily parasitized than genetically pure *M. musculus* or *M. domesticus*.

Excessive worminess seems not to be related to the sex of the mouse, as reported elsewhere[8]. Seven of the 15 mice with >500 nematodes, and three of the six animals most heavily infected with cestodes, were males. Also, the heavy infections are widely distributed within the hybrid zone and present during both collecting periods, suggesting that the causative agent does not come from some unique ecological or temporal factors in a particular mouse colony. No differences in habitat or mouse abundance were obvious between the collecting sites within the hybrid zone and in the adjacent regions. The 15 mice with the highest nematode burdens come from seven cities (Fig. 1), and the nine animals with heavy cestode infestations were collected at four locations (numbers 15, 18, 19 and 27).

Neither the high mean levels of nematodes in the hybrid zone ($\bar{x} = 198$) nor the extraordinarily high maximal counts (>500) found in some mice have been found in other sampled populations of wild *M. musculus*[9] or *M. domesticus*[10–14]. Although the pathological effects of these heavy worm burdens on wild mice have not been studied, laboratory experiments implicate pinworms in causing rectal prolapse, enteritis, liver granulomas and poor weight gains[15], as well as in altering immune responses to other disease agents[15]. Although the worm burdens are clearly not fatal to these wild mice, they seem likely to be deleterious.

The diversity of life cycles among these parasites is notable. The caecal pinworms and some of the intestinal *Hymenolepis* are transmitted directly from mouse to mouse but others have complex life cycles involving other animals as intermediate hosts. The epidemiological conditions needed to account for the worminess of these hybrids must therefore be diverse. The implication is that some of the mice are peculiarly susceptible to worm infections.

Because nearly all the wormy mice are hybrids, we suggest that the increased parasitism has a genetic basis. Reviews of the literature on resistance by small mammals to parasites[16–20] give many examples of genetically-based resistance to both uni- and multicellular parasites, including some of the helminths we found in the wild mice (*Aspiculuris*, *Hymenolepis*, *Trichuris*, and *Taenia*). The molecular and cellular mechanisms that mediate resistance to worms are diverse and are known to include immune-response and histocompatibility genes in mice[16,19,20]. It has been shown that the major genes for susceptibility to the bacterium *Salmonella typhimurium* reside on different chromosomes in *M. domesticus* and *M. musculus*[2], which implies different genetic mechanisms of resistance. Resistance to *Trichuris muris* and *Trichinella spiralis* behave as dominant characters. The reviews[16–20] emphasize the ubiquity and genetic simplicity of natural resistance.

These known cases of mouse resistance to parasites suggest that simple genetic models could explain our field observations. One possible model is that helminth resistance resides in different genes in *M. musculus* and *M. domesticus*, and that wormy mice have recombinant genotypes. If the resistance genes are dominant in the wild, as in the laboratory, then the susceptible, wormy mice would arise after the initial hybridization events, from a mating between recombinant animals. Alternatively, resistance could be conferred by a combination of alleles at two linked loci and the alleles conferring resistance could differ in the two species. Recombination would break up advantageous combinations of alleles. Parental and F_1 genotypes are rare in the hybrid zone; nearly all the mice have backcross genotypes[1]. Thus, these hybrid populations have genetic constitutions compatible with either of these genetic models.

The initial field observation (R.D.S.) which led to this quantitative analysis of helminth parasitism was the apparently greater numbers of fleas on the recently killed, hybrid mice, in comparison to non-hybrids. Conversations with colleagues who study hybridization in other species indicate that hybrids may frequently have more parasites than the adjacent parental populations (for example, gall aphids on cottonwoods, gill parasites in stone crabs, and intestinal worms in two genera of salamanders; T. Bert, T. Kocher, S. Sessions and T. Whitham, personal communication). If future studies verify higher-than-normal levels of parasitism and 'disease' in other hybrid zones, they would encourage the view that the loci which change fastest in response to selection after two populations separate are often those controlling resistance to parasites[21].

We thank I. Butler, K. Damme, J. Klein, R. Kraft, F. Seidl, H. Tichy, F. Pirchner and the people who allowed us to trap mice in their buildings, P. Bateson, W. Hamilton, F. James, A. Kuris, E. Mayr, E. Prager, J. Shields and S. Sweet for discussion about the manuscript, and the NSF for financial support.

Received 15 April; accepted 8 September 1986.

1. Sage, R. D., Whitney, J. B. III & Wilson, A. C. *Curr. Top. Microbiol. Immun.* **127**, 75–85 (1986).
2. O'Brien, A. D., Weinstein, D. L., D'Hoostelaere, L. D. & Potter, M. *Curr. Top. Microbiol. Immun.*, **127**, 309–318 (1986).
3. Barton, N. H. *Heredity* **44**, 277–278 (1980).
4. Kraft, R. *Saugetk. Mitt.* **32**, 1–12 (1985).
5. Zimmermann, K. *Zool. Jb. Abt. Syst.* (*Oekol.*) *Geogr. Tier.* **78**, 301–322 (1949).
6. Crofton, H. D. *Parasitology* **62**, 179–194 (1971).
7. Anderson, R. M. & Gordon, D. M. *Parasitology*, **85**, 373–398 (1982).
8. Sage, R. D. in *The Mouse in Biomedical Research* Vol. 1 (eds Foster, H. L., Small, J. D. & Fox, J. G.) 39–90 (Academic, New York, 1981).
9. Pleshchev, V. S. & Kozlov, A. N. *Mater. Nauchn. Konf. Vses. O. Gel'mint.* **30**, 135–139 (1978).
10. Elton, C. *Parasitology* **26**, 107–111 (1934).
11. Whitaker, J. O. *Proc. Indiana Acad. Sci.* **79**, 441–448 (1970).
12. Clark, D. E. *Proc. Indiana Acad. Sci.* **80**, 495–500 (1970).
13. Behnke, J. M. *J. Helminth.* **49**, 85–90 (1975).
14. Singleton, G. R. *Aust. Wildl. Res.* **12**, 437–445 (1985).
15. Wescott, R. B. in *The Mouse in Biomedical Research* Vol. 2 (eds Foster, H. L., Small, J. D. & Fox, J. G.) 373–384 (Academic, New York, 1982).

16. Wakelin, D. *Adv. Parasit.* **16,** 219-308 (1978).
17. Skamene, E., Kongshaun, P. A. & Landy, M. (eds) *Genetic Control of Natural Resistance to Infection and Malignancy* (Academic, New York, 1980).
18. Albright, J. F. & Albright, J. W. *Contemp. Top. Immunobiol.* **12,** 1-52 (1984).
19. Wassom, D. L., David, C. S. & Gleich, G. J. in *Genetic Control of Natural Resistance to Infection and Malignancy* (eds Skamene, E. Kongshavn, P. A. L. & Landy, M.) 75-82 (Academic, New York, 1980).
20. Wakelin, D. in *Ecology and Genetics of Host-Parasite Interactions* (eds Rollinson, D. & Anderson, R. M.) 39-54 (Academic, London, 1985).
21. Hamilton, W. D. in *Population Biology of Infectious Diseases* (eds Anderson, R. M. & May, R. M.) 269-296 (Springer, Berlin, 1982).

A single micromanipulated stem cell gives rise to multiple T-cell receptor gene rearrangements in the thymus *in vitro*

Gwyn T. Williams*, Rosetta Kingston*, Mike J. Owen†, Eric J. Jenkinson* & John J. T. Owen*

* Department of Anatomy, Medical School, Vincent Drive,
Edgbaston, Birmingham B15 2TJ, UK
† ICRF Tumour Immunology Unit, Zoology Department,
Gower Street, London WC1E 6BT, UK

The extensive range of specificities of T-cell receptors is generated, as for immunoglobulins, by rearrangement of genetic information[1]. Much valuable information about rearrangement processes has been inferred by comparing DNA from (monoclonal) lymphoid lines with germ-line DNA[2,3] and, for B cells, from rearrangements in some Abelson murine leukaemia virus-transformed cell lines[4,5]. However, because it is difficult to isolate and grow precursor populations, it has not proved possible to study rearrangements occurring in normal untransformed cells *in vitro*. Here we show that a single T-cell precursor colonizing an alymphoid thymus lobe in organ culture can generate multiple receptor β-chain gene rearrangements. These observations provide unequivocal evidence for the intra-thymic diversification of the T-cell repertoire. They also offer the possibility of investigating rearrangement and its control in the clonal progeny of a single normal T-cell precursor without the perturbations involved in the use of viral transformation or the production of T-cell hybridomas.

T cells require both specific antigen and self major histocompatibility (MHC) determinants[6,7] for activation. They possess a cell surface receptor made up of α- and β-chains[8,9] encoded by genes which are rearranged and expressed only in T lymphocytes[10,11]. The rearrangement of receptor β-chain genes during mouse fetal development has been monitored by hybridizing β-chain gene probes to restriction enzyme-digested DNA from cells at different stages of development[12,13]. The proportion of thymic lymphocytes showing rearrangement of the β-chain genes is very small at day 14, then increases progressively. Several interpretations of these observations are possible: (1) gene rearrangement may occur in the thymus; (2) rearranged populations may proliferate preferentially in the thymus; (3) there may be progressive recruitment of rearranged cells into the thymus; or some combination of these processes. To distinguish between these possibilities we used a recently developed technique[14] to recolonize an alymphoid thymic lobe with a single blast cell derived from 14-day fetal thymus. Successfully recolonized lobes yield up to 1.1×10^5 lymphocytes, and contain various T-cell marker defined sub-populations, all derived from the single donor cell[14]. If rearrangement had already occurred at the 14-day stage, either within the thymus or at a pre-thymic level, one or two rearranged bands would be detected. If no rearrangement had occurred by the 14-day stage or during organ culture a strong germ-line pattern would still be evident, while

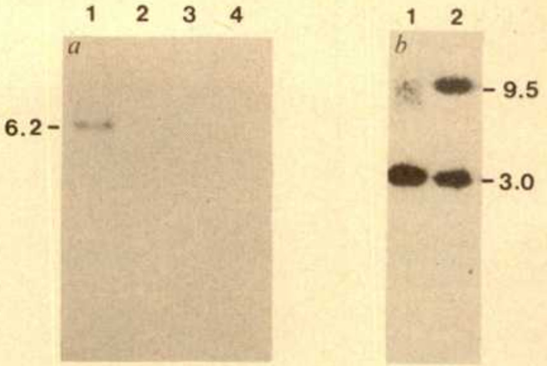


Fig. 1 Southern blot analyses of T-cell receptor β-chain genes in the populations produced in thymus lobes recolonized by a single T-cell precursor. *a*, *Pvu*II digestion of: lane 1, liver DNA; lanes 2-4, progeny of single cells; *b*, *Hin*dIII digestion of: lane 1, progeny of single cell; lane 2, liver DNA. Alymphoid CBA (H-2^k) thymus lobes were prepared and each was recolonized by a single T-cell precursor taken from a 14 day of gestation BALB/c thymus lobe as described previously[14]. *a*, DNA was prepared from the lymphoid cells produced from single precursors, digested with *Pvu*II, separated on an agarose gel, Southern blotted and hybridized with a ^{32}P-labelled probe for the constant region of the β-chain gene. Filters were washed at high stringency before autoradiography. Liver DNA (β-chain gene complex in germ-line state) was analysed in parallel. The two constant region genes are not resolved and migrate with an apparent size of 6.2 kb (liver DNA, lane 1). The substantial reduction in these bands in the progeny of single cells, with no dominant rearranged band, indicates that multiple rearrangements have occurred (lanes 2-4). Five other lobes analysed all showed similar results (data not shown). The filter was rehybridized with a probe for β_2-microglobulin, confirming that about the same amount of DNA was present in all tracks. *b*, DNA from the progeny of a single cell (lane 1) and liver DNA (lane 2) were digested with *Hin*dIII and analysed as above. In the germ-line state (lane 2) the $C_\beta 1$ restriction fragment migrates with an apparent size of 9.5 kb, and the $C_\beta 2$ fragment with an apparent size of 3.0 kb. In the DNA from the single cell progeny the germ-line $C_\beta 1$ fragment has been sharply reduced and a smear of bands has appeared. This indicates that heterogeneous multiple rearrangements have occurred within the thymus.

Methods. Alymphoid lobes were obtained after culture in the presence of deoxyguanosine (dGuo) and associated with single micromanipulated precursors for 24 h in hanging drops formed in Terasaki plates, with subsequent transfer to organ culture for 11-13 days for optimal expansion of successfully recolonized lobes. This produces lobes exclusively recolonized by progeny of the single donor precursor[14]. At the end of the organ culture period lymphoid cells were released from individual thymus lobes by gentle teasing with fine cataract knives and the resultant suspension was washed in Tris-buffered Earles saline and frozen at −80 °C before preparation of DNA. For *Pvu*II digestion DNA was prepared by proteinase K digestion in 0.5% SDS followed by phenol-chloroform extraction and extensive dialysis. After *Pvu*II digestion to completion, the DNA was precipitated with ethanol and separated by electrophoresis on a 0.8% agarose gel. For *Hin*dIII digestion the frozen cell pellet was digested with proteinase K in 0.5% sodium lauroyl sarkosinate and 1% low gelation temperature agarose (Bio-Rad). The agarose was allowed to solidify at 0 °C and detergent was removed by incubation with several changes of 10 mM Tris 1 mM EDTA *p*H 8.0 (G.T.W., manuscript in preparation). The DNA was digested with *Hin*dIII in the presence of 0.1% nuclease-free bovine serum albumin and separated on a 0.6% agarose gel (normal gelation temperature). Gels were blotted onto Zeta-probe (Bio-Rad) membranes and hybridized with an oligo-labelled[19] C_β probe derived from plasmid 86T1[11]. Filters were washed at high stringency and autoradiographed for 10 days using Kodak XAR-5 film and two intensifying screens (Dupont Cronex Lightning-plus).

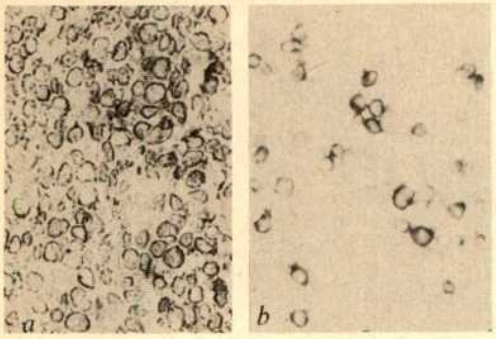


Fig. 2 Detection of T-cell receptor β-chain expression in the progeny of a single T-cell precursor. *a*, Labelling with anti-Lyt 2.2; *b*, labelling with F-23.1, a monoclonal antibody that recognizes the products of a V_β gene family present on approximately 25% of peripheral T cells[15]. Alymphoid CBA (H-2^k, Lyt 2.1) thymus lobes were recolonized by a single BALB/c (H-2^d, Lyt 2.2) precursor (Fig. 1 legend). After 11–13 days of organ culture individual lobes were snap frozen and four frozen sections prepared for indirect immunoperoxidase labelling. Two sections taken from the same lobe are shown. *a*, Labelling with monoclonal anti-Lyt 2.2 (NEN); most lymphoid cells are labelled, indicating that they have undergone differentiation in terms of T-cell marker expression and that they are derived from the single BALB/c (Lyt 2.2) precursor; as shown in our previous work, dGuo-treated lobes recolonized by a single cell contain T cells exclusively of donor type[13,20]. *b*, The labelling with F-23.1 demonstrates productive rearrangement in a proportion of cells, but does not provide a quantitative estimate of the extent of T-cell receptor β-chain diversity, as the proportion of receptor-positive cells is unknown. However, the observed proportion of F-23.1^+ cells is fully consistent with the heterogeneous multiple β-chain gene rearrangements shown in Fig. 1. The large reduction in the C_β germ-line bands shown in Fig. 1 suggests that most of the F-23.1^- cells have also rearranged the β-chain gene locus, although not all of these rearrangements may be productive and a proportion of cells may not synthesize any β-chain. The F-23.1 labelling might detect surface and/or cytoplasmic β-chain expression[21] (in a recent study using a double immunofluorescence procedure we have found that, in newborn thymus, approximately equal proportions of cytoplasmic F-23.1^+ and surface F-23.1^+ cells are present[22]). As further proof that the progeny of a single cell can proceed to productive receptor expression, alymphoid CBA lobes (Thy 1.2) were recolonized by single AKR (Thy 1.1) precursors. Double immunofluorescence labelling with anti-Thy-1.1 and F-23.1 showed that 90% of cells were Thy-1.1^+ and 3.3% of these were surface F-23.1^+, a value similar to that observed in non-manipulated thymus organ cultures[13].

multiple rearrangements in the progeny of the single cell in the organ-cultured thymus would reduce the germ-line bands without producing any dominant rearranged bands.

The lymphocyte DNA from several lobes, each recolonized by a single cell, was prepared and analysed separately by restriction digestion, gel electrophoresis, Southern blotting and hybridization with radiolabelled β-chain constant region probes (Fig. 1). For *Pvu*II-digested DNA the intensity of the 6.2-kilobase (kb) band corresponding to the germ-line ($C_\beta1$ and $C_\beta2$) genes decreased markedly in lymphocytes from recolonized lobes, indicating extensive multiple rearrangements (Fig. 1*a*). In *Hin*dIII-digested liver DNA a 9.5-kb band corresponds to $C_\beta1$ and a 3.0-kb band to $C_\beta2$. Rearrangement to either β-chain constant region reduces the intensity of the 9.5-kb band through deletion or rearrangement of $C_\beta1$, but does not affect the 3.0-kb fragment, which thus acts as an internal marker[11,12]. The 9.5-kb ($C_\beta1$) band is sharply reduced in the DNA of the progeny of the single precursor showing that the cells have undergone multiple β-chain gene rearrangements (Fig. 1*b*). The heterogeneity of the rearrangements in the lymphoid cells is confirmed by the appearance of a smear of rearranged bands. All nine lobes analysed showed multiple rearrangements.

To demonstrate productive β-chain gene rearrangements from a single thymic stem cell we used a monoclonal antibody (F-23.1), which reacts with a sub-population of β-chain variable regions[15] to look for expression of β-chain. A significant proportion of the cells produced from single precursors express T-cell receptor β-chains (Fig. 2), so many of the gene rearrangements detected are indeed productive.

The recolonization of alymphoid thymic lobes by single micromanipulated precursors has allowed us to provide a definitive demonstration of gene rearrangement involved in the generation of T-cell receptor β-chain diversity in the thymus. Neither progressive recruitment nor preferential proliferation of cells with rearranged β-chain genes could account for these observations as neither is possible in this system. This is the first *in vitro* demonstration of generation of diversity in any receptor molecule of the immune system from a single untransformed cell. Since these results show that the thymic environment is sufficient to induce β-chain gene rearrangements in T-cell precursors, they argue against the need for pre-thymic rearrangement as suggested by some functional studies[16,17]. Furthermore, 14-day thymus cells, although capable of proliferation *in vitro*, do not undergo rearrangement outside the thymus[18]. Thus the thymic environment is important in initiating the rearrangement process. This system provides a method for investigating intercellular interactions in the thymus, as well as the molecular mechanisms which produce gene rearrangements during T-cell differentiation. The technique should also prove invaluable in identifying pre-thymic T-cell precursors in normal and transformed populations by their capacity to undergo receptor gene rearrangement when introduced into a thymus. The pattern of rearrangement and β-chain utilization obtained when progressively more mature cells are re-introduced into the thymus may also be analysed using this approach.

This work was supported by the MRC (E.J.J., R.K. and J.J.T.O.), the Rowbotham Bequest (G.T.W., E.J.J) and Wellcome Trust (G.T.W.). We thank Dr M. M. Davies for clone 86T1, Dr M. Bevan for monoclonal antibody F23.1, Dr K. Docherty, Mr I. Phillips and Dr F. H. C. Franklin for help and advice, Alison Orchard for assistance with the photography and Claire Hundley for typing the manuscript.

Received 24 April; accepted 28 August 1986.

1. Hood, L., Kronenberg, M. & Hunkapiller, T. *Cell* **40,** 225–229 (1985).
2. Tonegawa, S. *Nature* **302,** 575–581 (1983).
3. Kronenberg, M., Siu, G., Hood, L. & Shastri, N. *A. Rev. Immun.* **4,** 529–591 (1986).
4. Lewis, S., Rosenberg, N., Alt., F. & Baltimore, D. *Cell* **30,** 807–816 (1982).
5. Sugiyama, H. *et al. Nature* **303,** 812–815 (1983).
6. Zinkernagel, R. M. & Doherty, P. C. *Nature* **248,** 701–702 (1974).
7. Seidman, J. G. & Leder, P. *Nature* **276,** 790–795 (1978).
8. Allison, J., McIntyre, B. & Block, D. *J. Immun.* **129,** 2293–2300 (1982).
9. Haskins, K. *et al. J. exp. Med.* **157,** 1149–1169 (1983).
10. Yanagi, Y. *et al. Nature* **308,** 145–149 (1984).
11. Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308,** 149–153 (1984).
12. Snodgrass, H. R., Kisielow, P., Kiefer, M., Steinmetz, M. & von Boehmer, H. *Nature* **313,** 592–595 (1985).
13. Owen, M. J., Jenkinson, E. J., Williams, G. T., Kingston, R. & Owen, J. J. T. *Eur. J. Immun.* **16,** 875–878 (1986).
14. Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **317,** 811–813 (1985).
15. Staerz, U. D., Rammensee, H. G., Benedetto, J. & Bevan, M. J. *J. Immun.* **134,** 3994–4000 (1985).
16. Morisey, P. J., Kruisbeck, A. M., Shanan, S. O. & Singh, H. *Proc. natn. Acad. Sci. U.S.A.* **79,** 2003–2007 (1982).
17. Chervenak, R., Cohen, J. J. & Miller, S. D. *J. Immun.* **131,** 1688–1692 (1983).
18. Palacious, R. & von Boehmer, H. *Eur. J. Immun.* **16,** 12–19 (1986).
19. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132,** 6–13 (1983).
20. Jenkinson, E. J., Franchi, L. L., Kingston, R. & Owen, J. J. T. *Eur. J. Immun.,* **12,** 583–587 (1982).
21. Farr, A. G., Anderson, S. K., Marrack, P. & Kappler, J. *Cell* **43,** 543–550.
22. Owen, J. J. T., Kingston, R. & Jenkinson, E. J. *Immunology* **59,** 23–27 (1986).

Extensive proliferation of mature connective-tissue type mast cells *in vitro*

Tatsutoshi Nakahata*, Toshimi Kobayashi*, Akira Ishiguro*, Kohichiro Tsuji*, Kuniaki Naganuma*, Osaaki Ando*, Yoshio Yagi*, Kenji Tadokoro† & Taro Akabane*

* Department of Pediatrics, Shinshu University School of Medicine, Asahi 3-1-1, Matsumoto 390, Japan
† Department of Medicine and Physical Therapy, University of Tokyo, Tokyo 113, Japan

There are two phenotypically distinct subpopulations of mast cells in rodents: connective tissue-type mast cells (CTMC) and mucosal mast cells (MMC). These populations differ in their location, cell size, staining characteristics, ultrastructure, mediator content and T-cell dependency[1–3]. Several investigators[4–9] recently reported a further subclass of mast cells which arise when normal mouse haematopoietic cells are cultured with interleukin-3 (IL-3); IL-3 is an activity similar or identical to mast-cell growth factor, histamine-producing factor, or P-cell stimulating factor[4–9]. These cultured mast cells are in many ways similar to MMC; they stain with Alcian blue but not safranin, contain chondroitin sulphate E proteoglycan rather than heparin proteoglycan and have relatively low histamine content, as do MMC[10–13]. Although proliferation of MMC is known to be T-cell dependent *in vivo* and thought to be IL-3-dependent *in vitro*, the factors on which CTMC proliferation depends remain elusive. Here we show that mature CTMC purified from mouse peritoneal cells can proliferate *in vitro* in methylcellulose culture and maintain the appearance and function of CTMC. We also present evidence that mature CTMC cannot proliferate in the presence of pure IL-3 alone.

When normal mouse peritoneal cells were cultured in methylcellulose medium containing pokeweed mitogen-stimulated spleen-cell-conditioned medium (PWMSCM), pure mast-cell colonies developed[14]. Detailed histochemical analyses of individual mast-cell colonies, together with observations of colony shape and size, allowed us to classify them into two types, type 1 and type 2 (Fig. 1*a*,*d*). The primary difference between colony types 1 and 2 was size: type 1 were smaller and more compact and revealed no signs of degeneration over 30 days of cultivation without additional feeding. Type 1 colonies contained 20 to 500 large mast cells with a single nucleus and many large granules (Fig. 1*b*); type 2 usually contained >2,000 small mast cells, with relatively small granules on day 16 of cultivation (Fig. 1*e*) and occasionally bilobed nuclei.

The two types of colonies could be more clearly distinguished by staining with Alcian blue–safranin or Berberine sulphate, or by measuring cellular histamine content. Cytocentrifuge preparations stained with Alcian blue–safranin and Berberine sulphate showed that most of the type 1 mast cells were both safranin-positive and Berbine sulphate-positive, whereas most of the type 2 mast cells stained with Alcian blue but not with safranin or Berberine sulphate (Table 1). Berberine sulphate staining has been reported to be highly specific for connective tissue-type mast cells, which contain heparin[15,16]. We confirmed this by demonstrating that the staining was blocked by treating the cells with heparinase but not chondroitinase ABC. The histamine content of type 1 cells was significantly greater than that of type 2 cells (Table 1). These data strongly suggest that most type 1 cells are CTMC-like[1–3,17]; by contrast, the type 2 cells resemble bone marrow-derived cultured mast cells, which have some characteristics in common with MMC[3,10,11,17].

We separated the precursors of type 1 mast-cell colonies from those of type 2 mast-cell colonies by removal of phagocytes followed by density gradient centrifugation (Table 2). Both types of precursor was enriched by the removal of phagocytes. Eight fractions were separated by Percoll density-gradient

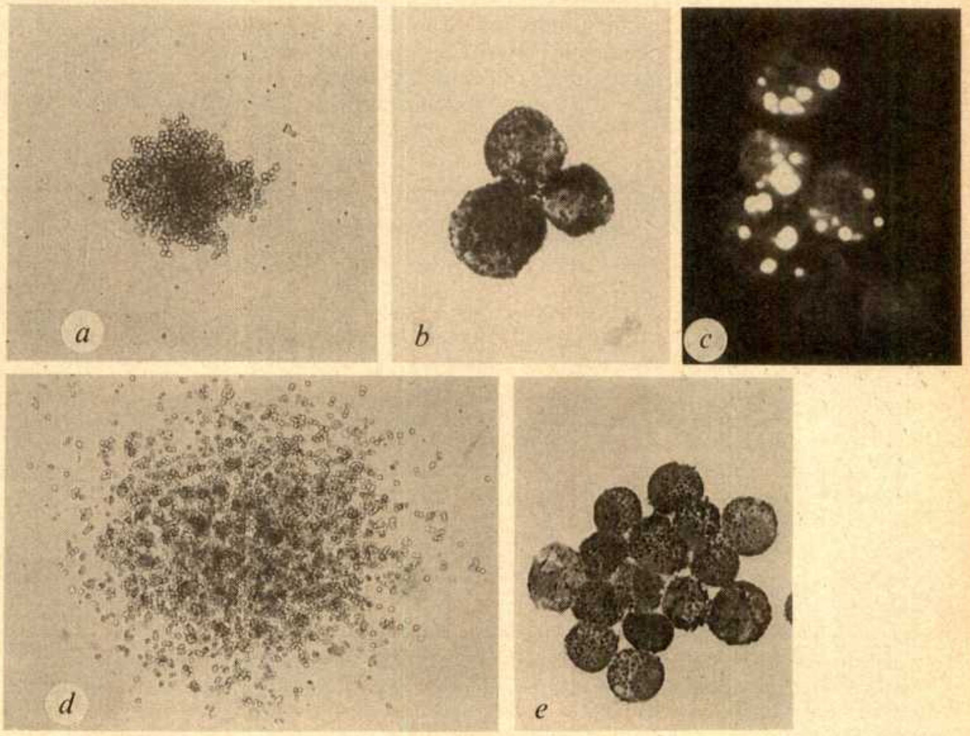


Fig. 1 Mast-cell colonies. *a*, Typical type 1 mast-cell colony on day 16 of culture; *b*, May-Grunwald–Giemsa-stained smear on the type 1 mast-cell colony shown in *a*; *c*, Berberine sulphate staining of the same colony; most mast cells show fluorescence; *d*, typical type 2 mast-cell colony on day 16 of culture; *e*, May-Grunwald–Giemsa-stained smear of the type 2 mast-cell colony shown in *d*.

Methods. Peritoneal cells from 20–25-week-old BDF$_1$ mice were collected as described previously[14]. Methylcellulose cultures were carried out by using a modification of the technique described previously[14,29]. Culture mixture (1 ml) containing 1×10^5 peritoneal cells, α-medium (Flow), 0.8% methylcellulose (Shinetsu Kagaku Co), 30% fetal bovine serum (Flow), 1% deionized bovine serum albumin (Sigma), 10^{-4} M mercaptoethanol (Eastman) and serum-free PWMSCM was incubated at 37 °C in a humidified atmosphere flushed with 5% CO_2 in air. Serum-free PWMSCM was prepared as described previously[30]. Mast-cell colonies containing ≥20 cells were counted on day 16. Individual colonies were lifted from methylcellulose medium using 3 μl Eppendorf pipette under a microscope and collected in Eppendorf microcentrifuge tubes containing 0.5 ml phosphate-buffered saline. After washing, the samples were spun in a cytocentrifuge (Shandon) (800 r.p.m. for 5 min) and stained with May-Grunwald–Giemsa, Alcian blue–safranin[31] or Berberine sulphate[16].

Table 1 Differences between types 1 and 2 mast-cell colonies

Type of mast-cell colony	No. of cells per colony	Cell size	Positive cells stained with Berberine sulphate	Alcian blue	Safranin	Histamine content (ng per 10^5 cells)
1	258 ± 35*	18.9 ± 1.9 μm	65.1 ± 21.2	77.4 ± 8.8	98.8 ± 0.9	64.74
	(20–500)†	(15–23)*	(18–98)	(55–87)	(94–100)	
2	8,500 ± 558	13.3 ± 3.7	0	100	1.4 ± 0.4	5.88
	(2,000–18,500)	(7–21)			(0–5)	

Colony size was determined by direct cell counting *in situ* with ×200 magnification on an inverted microscope when colonies consisted of less than 500 cells. When colonies contained >500 cells, a counting chamber was used. We analysed 50 types 1 and 2 mast-cell colonies. The histamine content was assayed using pooled colonies.

* Mean ± s.d.; † range.

centrifugation and individual fractions were analysed for cellular composition and relative concentration of the two types of precursor. The proportion of morphologically identifiable mast cells was increased in the heavier fractions and roughly paralleled the number of type 1 colonies. For example, fraction 7 (≥ 1.13 g ml^{-1}) was an almost pure population of large mast cells all of which stained with both safranin and Berberine sulphate but not with Alcian blue, and were thus classed as mature CTMC[18]. When these cells were cultured in methylcellulose medium containing PWMSCM approximately half the cells gave type 1 colonies, but no type 2 colonies developed (experiment 2, Table 2). It is of interest that, although these colonies contained mostly safranin-positive, Berberine sulphate-positive cells, about half the colonies contained small numbers of Alcian blue-positive, safranin-negative cells which are thought to be a more primitive stage than CTMC[18]. The clonal origin of type 1 colonies was confirmed using single-cell culture[19]. Cells in fraction 7 were plated at a concentration of 1×10^3 per ml in methylcellulose medium containing PWMSCM in 35-mm plastic dishes. A single mast cell was identified morphologically, lifted from the medium using a fine Pasteur pipette attached to a Leitz micromanipulator, transferred to a fresh culture dish and incubated for 16 days. Fifteen of 50 transferred single mast cells produced colonies, all of which were type 1. Thus some morphologically identifiable mature CTMC can proliferate *in vitro* to give colonies containing cells with CTMC-like features in the absence of accessory cells. This conclusion is supported in part by recent *in vivo* studies indicating that some peritoneal mast cells retain extensive proliferative potential after injection into the skin of mast-cell-deficient mice[20].

Sequential replating experiments of individual type 1 colonies derived from the cells in fraction 7 confirmed the proliferation potential of mature CTMC. Most of the primary colonies produced many secondary and tertiary colonies in methylcellulose medium containing PWMSCM. Although all the primary colonies contained mostly Berberine sulphate-positive, safranin-positive cells, there was a gradual decrease in the incidence of Berberine sulphate-positive mast cells and an increase in that of Alcian blue-positive, safranin-negative mast cells in individual secondary and tertiary colonies (data not shown). These findings suggest that some mature CTMC may be able to 'regress' and become a more immature type of CTMC or acquire some of the features of MMC. It will be necessary to perform a detailed biochemical examination of individual colonies to confirm this possibility.

Cells in fraction 2 (≤ 1.05 g ml^{-1}; lymphocytes, lymphoblastoid cells and no detectable mast cells) produced only type 2 mast-cell colonies, consisting of Alcian blue-positive cells which did not stain with either safranin or Berberine sulphate (Table 2). The number of type 2 mast-cell colonies that developed was linearly related to the number of cells plated (data not shown), suggesting that the colonies were of clonal origin. Type 2 precursors of mast-cell colonies have a similar (low) density to bone marrow precursors of factor-dependent cultured mast cells[21].

Table 2 Cellular composition and concentration of mast-cell precursors in fractionated peritoneal cells

Experiment	Origin of cells	Cellular composition: Mast cells	Lymphoid cells	Number of colonies per 10^5 cells: Type 1 mast cell colony	Type 2 mast cell colony	Total plating efficiency
1	Unseparated cells	0.9%	53.9%	102 ± 25*	2 ± 1	0.1%
	Nonphagocytic cells	1.6	98.1	308 ± 15	4 ± 1	0.3
	Fraction 2	0	100	0	22 ± 2	0.02
	3	0	100	0	7 ± 2	0.01
	4	1.8	98.2	321 ± 28	3 ± 1	0.3
	5	97.9	2.1	8,230 ± 151	0	8.2
	6	100	0	19,930 ± 505	0	19.9
	7	100	0	16,210 ± 309	0	16.2
2	Fraction 2	0	100	0	12 ± 1	0.01
	7	100	0	50,608 ± 2,131	0	50.6

Phagocytic cells were removed using carbonyl iron[32]. Percoll (Pharmacia) was made isosmotic by the addition of 1 part of 10× Hanks' solution containing 1% bovine serum albumin and 200 mM HEPES to 9 parts of Percoll. Continuous density gradients of Percoll were generated by centrifugation of 7 ml of Percoll solution at 23,000g for 60 min at 4 °C in a Tomy RS-20 FIIIS fixed rotor. Cell suspension (1 ml) containing 1×10^7 nonphagocytic cells was layered on the top of Percoll continuous gradients and centrifuged at 500g for 30 min[20]. Although eight fractions were separated, fractions 1 and 8 were discarded because the former contained degenerated cells and the latter contained a small number of macrophages which phagocytosed iron particles. Cellular composition was determined on cytocentrifuge preparations stained with May-Grunwald-Giemsa. Each fraction was cultured before (Fig. 1 legend).

* Mean ± s.d. of triplicated cultures.

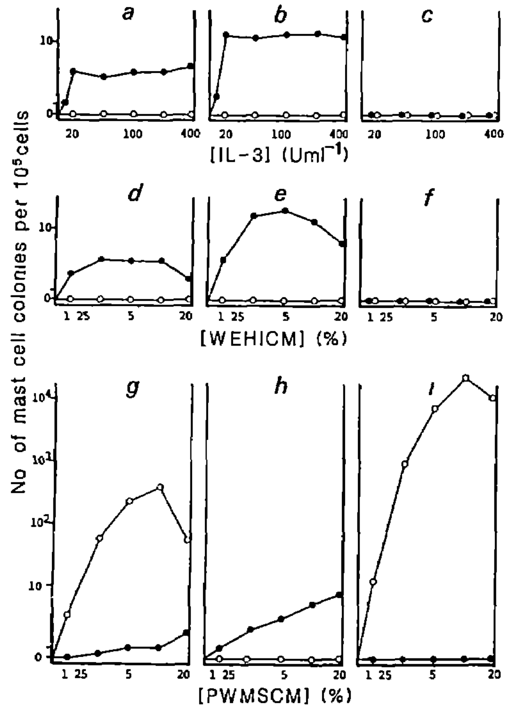


Fig. 2 Dependence of colony growth on growth factors Growth of colony types 1 and 2 from *a*, *d*, *g*, unseparated peritoneal cells, *b*, *e*, *h*, cells in fraction 2, *c*, *f*, *i*, cells in fraction 7 at varying concentrations of *a*–*c*, pure IL-3 in culture, *d*–*f*, WEHICM, and *g*–*i*, PWMSCM Open circles, type 1 mast-cell colony, filled circles, type 2 colony

Thus the precursors for two distinct types of mast cell can be separated by Percoll density-gradient centrifugation

We investigated the growth factor requirements of the two types of mast-cell precursors Homogeneous populations of mast cells resembling MMC are obtained *in vitro* from mouse bone marrow, spleen, fetal liver and lymphoid tissue in cultures supplemented with conditioned medium from stimulated T cells or the WEHI-3B cell line[4-7] It has been demonstrated that the biologically active component in the conditioned medium is IL-3[8,9,12,22] However, little is known about the factor(s) supporting CTMC proliferation *in vitro* Complex culture systems have been reported in which heparin-containing mast cells differentiate when the culture contains embryonic fibroblasts and lymphocytes from mice preimmunized to horse serum[23] Recent observations show that rat CTMC can be maintained *in vitro* in the presence of living fibroblasts but do not divide, and that fibroblast-derived conditioned medium does not have any ability to maintain CTMC[24] We next examined whether conditioned media (other than PWMSCM) or pure IL-3 support the proliferation of either type 1 or type 2 colonies in our methylcellulose culture system Neither type of colony will grow in the presence of lung-conditioned medium (containing granulocyte-macrophage colony stimulating factor and granulocyte-colony stimulating factor[25,26]) L-cell conditioned medium (containing macrophage colony stimulating factor[27]) also failed to stimulate colony growth

Figure 2 shows the plating efficiencies of two types of mast-cell colonies from unfractionated or fractionated peritoneal cells stimulated by different concentrations of pure IL-3, WEHI-3B conditioned medium (WEHICM) or PWMSCM No type 1 mast-cell colonies were observed in the presence of up to 400 U ml^{-1} of pure IL-3 in the culture WEHICM also failed to stimulate type 1 mast-cell colony growth PWMSCM stimulated type 1 mast-cell colony growth from unfractionated cells and fraction 7 significantly at a maximum concentration of 10% Pure IL-3 produced type 2 mast-cell colonies from unfractionated cells and fraction 2 in a dose-dependent fashion, reaching a plateau at 20 U ml^{-1} IL-3 Pure IL-3 also supported the growth of secondary mast-cell colonies consisting of cells which stained with Alcian blue but not with safranin or Berberine sulphate upon replating of type 2 colonies Our experiments show that the two types of mast-cell precursors are different in origin and in requirement for growth factors The IL-3 dependent growth of the low density precursors for type 2 colonies is consistent with a previous report that bone marrow precursors are distinct from mast cells, have a low density and proliferate in response to IL-3[21] Our PWMSCM contained some factor(s) that stimulate mature CTMC proliferation and is not IL-3 However, our data do not exclude the possibility that IL-3 can act as a co-stimulator for CTMC proliferation, because there was a small amount of IL-3 in the cultures containing 10% (v/v) PWMSCM (3 6 U ml^{-1} 100 U activity was arbitrarily defined as that amount of IL-3 per ml that gives 30% of maximal ^{3}H-thymidine incorporation into the IL-3 dependent cell line 32 Dcl[28]) Further studies are needed to determine the nature and source of the factor(s) responsible for mature CTMC proliferation

The CTMC culture system described here should allow the elucidation of the cellular and molecular mechanisms that control the proliferation of mature CTMC An examination of the changes in phenotype and processes involved in turnover of mediators during extensive proliferation should also be possible

We thank Drs Y Kitamura and A Komiyama for helpful advice and Dr J N Ihle for pure IL-3 Supported by grants from the Ministry of Education, Science and Culture, Japan

Received 5 March, accepted 6 August 1986

1 Enerback, L *Monogr Allerg* **17,** 222–232 (1981)
2 Bienenstock, J, Befus, A D, Pearce, F, Denburg, J & Goodacre, R *Allerg clin Immun* **70,** 407–412 (1982)
3 Galli, S J, Dvorak, A M & Dvorak, H F *Prog Allerg* **34,** 1–141 (1984)
4 Nabel, G, Galli, S J, Dvorak, A M, Dvorak, H F & Cantor, H *Nature* **291,** 332–334 (1981)
5 Schrader, J W, Lewis, S J, Clarke-Lewis, I & Culvenor, J G *Proc natn Acad Sci USA* **78,** 323–327 (1981)
6 Razin, E, Cordon Cardo, C & Good, R. A *Proc natn Acad Sci USA* **78,** 2559–2561 (1981)
7 Yung, Y P, Eger, R, Tertian, G & Moore, M A S *J Immun* **127,** 794–799 (1981)
8 Ihle, J N *et al. J Immun* **131,** 282–287 (1983)
9 Yokota, T *et al. Proc natn Acad Sci USA* **81,** 1070–1074 (1984)
10 Haig, D M, McKee, T A, Jarrett E E, Woodbury, R & Miller, H R P *Nature* **300,** 188–190 (1982)
11 Razin, E, Stevens, R. L, Akiyama, F, Schmidt, K & Austen, K *J biol Chem* **257,** 7229–7236 (1982)
12 Razin, E, *et al. J Immun.* **132,** 1479–1486 (1984)
13 Galli, S J *et al. J Cell Biol* **95,** 435–444 (1982)
14 Kobayashi, T *et al. J Immun* **136,** 1378–1384 (1986)
15 Enerback, L *Histochemistry* **42,** 301–313 (1974)
16 Dimlich, R V M, Meineke, H A, Reilly, F D & McCuskey, R S *Stain Tech.* **55,** 217–223 (1980)
17 Nakano, T *et al J exp Med.* **162,** 1025–1043 (1985)
18 Pretlow, T G & Gassady, I M *Am J Path* **61,** 323–338 (1970)
19 Nakahata, T *et al. Blood* **65,** 1010–1016 (1985)
20 Sonoda, T *et al. J exp Med* **160,** 138–151 (1984)
21 Yung, Y P, Wang, S Y & Moore, M A S *J Immun.* **130,** 2843–2848 (1983)
22 Rennick, D M *et al. J Immun* **134,** 910–914 (1985)
23 Bland, C E, Ginsburg, H, Silbert, J E & Metcalfe, D D *J biol Chem* **257,** 8661–8666 (1982)
24 Levi Schaffer, F *et al. J Immun* **135,** 3454–3462 (1985)
25 Burgess, A W, Camakaris, J & Metcalf, D *J biol Chem.* **252,** 1998–2003 (1977)
26 Nicola, N A, Metcalf, D, Matsumoto, M & Johnson, G R *J biol Chem* **258,** 9017–9023 (1983)
27 Stanley, E R. & Heard, B M *J biol Chem.* **252,** 4305–4312 (1977)
28 Frei, K, Bodmer, S, Schwerdel, C & Fontana, A *J Immun* **135,** 4044–4047 (1985)
29 Nakahata, T, Spicer, S S, Cantey, J R & Ogawa, M *Blood* **60,** 352–361 (1982)
30 Nakahata, T & Ogawa, M *Proc natn Acad Sci USA* **79,** 3843–3847 (1982)
31 Spicer, S S *J Histochem Cytochem* **8,** 18–35 (1960)
32 Kubota, K, Mizoguchi, H, Miura, Y, Suda, T & Takaku, F *Expl Hemat* **8,** 339–344 (1980)

A chimaeric receptor allows insulin to stimulate tyrosine kinase activity of epidermal growth factor receptor

Heimo Riedel, Thomas J. Dull, Joseph Schlessinger* & Axel Ullrich

Department of Developmental Biology, Genentech, Inc ,
460 Point San Bruno Boulevard, South San Francisco,
California 94080, USA

The cell surface receptors for insulin and epidermal growth factor (EGF) appear to share a common evolutionary origin, as suggested by structural similarity of cysteine-rich regions in their extracellular domains and a highly conserved tyrosine-specific protein kinase domain[1-3]. Only minor similarity is found outside this catalytic domain, as expected for receptors that have different ligand specificities and generate different biological signals[4,5]. The EGF receptor is a single polypeptide chain[6] but the insulin receptor consists of distinct α and β subunits[7] that function as an $\alpha_2\beta_2$ heterotetrameric receptor complex[8-11]. Provoked by this major structural difference in two receptors that carry out parallel functions, we have designed a chimaeric receptor molecule comprising the extracellular portion of the insulin receptor joined to the transmembrane and intracellular domains of the EGF receptor to investigate whether one ligand will activate the tyrosine kinase domain of the receptor for the other ligand. We show here that the EGF receptor kinase domain of the chimaeric protein, expressed transiently in simian cells, is activated by insulin binding. This strongly suggests that insulin and EGF receptors employ closely related or identical mechanisms for signal transduction across the plasma membrane.

Cloned human EGF receptor[1] and insulin receptor complementary DNAs[2,3] were used in combination with simian virus 40 (SV40) early promoter transcriptional elements (Fig 1*a*) to generate expression constructs coding for the complete human insulin receptor (HIR) and a chimaeric receptor (IER), shown in Fig 1*b* The insulin receptor is normally made as a precursor and proteolytically processed to generate the α and β subunits[4] IER was generated by fusing the coding region of the signal sequence and extracellular portion of the insulin receptor precursor (amino acids −27–914) to the coding sequence of the transmembrane and cytoplasmic domains of the EGF receptor[1] (amino acids 622–1,186) Using this chimaeric construct we asked whether the hybrid receptor would be proteolytically processed into insulin receptor α subunit and the recombinant IβE subunit, whether a functional complex would be formed and directed to the plasma membrane and whether insulin binding could stimulate autophosphorylation of the EGF receptor cytoplasmic domain in the chimaeric cell surface membrane glycoprotein

Expression of the complete insulin receptor resulted in protein products of relative molecular mass (M_r) 95,000 (95 K) and 180 K and a diffuse band around 135 K, as determined by ^{35}S-methionine metabolic labelling (Fig 2*a*) These polypeptides represent the β-subunit, the uncleaved precursor and the glycosylated α subunit of the insulin receptor[2,9,12-14] Proteolytic processing of the native insulin receptor precursor is inefficient under transient expression conditions in COS-7 cells, as one can see from the amount of HIR precursor remaining

By analogy with the native receptors, COS-7 cells expressing the IER receptor chimaera would be expected to synthesize a 135 K glycosylated α-chain, a comparably sized 120 K glycopolypeptide composed of the intracellular and transmem-

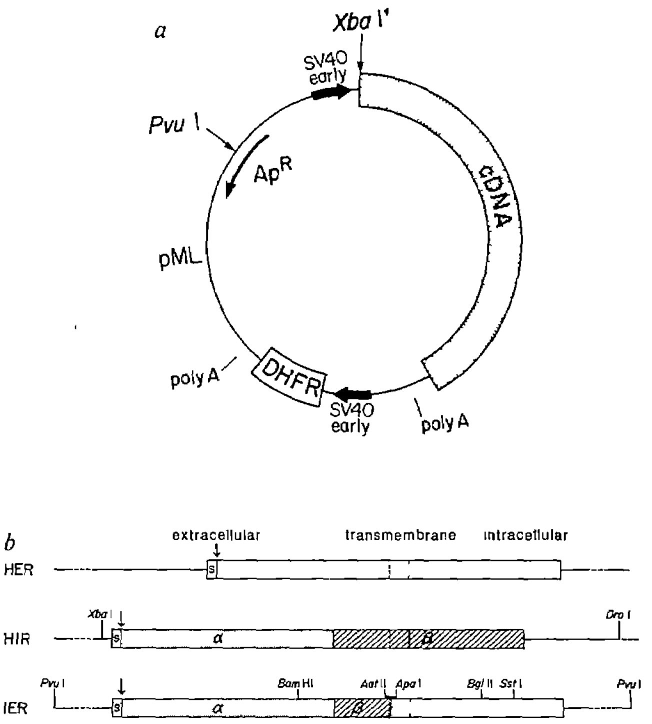


Fig. 1 *a*, Receptor expression plasmid, *b*, schematic comparison of normal and chimaeric receptor constructs *a*, The region encoding different receptor constructs is shaded (cDNA) Early SV40 promoter sequences are shown by heavy black arrows and poly-(A) addition sites are indicated The dihydrofolate reductase (DHFR) coding sequence[19] and restriction sites used for plasmid constructions are shown Expression of both cDNAs was controlled by promoter sequences of the SV 40 early region Sequences of the *Escherischia coli* plasmid pML (ref 20) were present to allow plasmid replication in *E coli* *b*, Coding regions of human EGF receptor (HER), human insulin receptor (HIR) and insulin-EGF receptor chimaera (IER) Shaded regions, HIRα subunit sequences, striped regions, HIRβ sequences, unshaded regions, HER sequences The sequences are aligned at their transmembrane domains (not to scale) S, signal sequences, arrows, signal sequence cleavage sites, vertical lines, receptor precursor cleavage sites, zigzag line, junction between heterologous receptor cDNAs, black bar, synthetic oligonucleotide used for the connection

Methods. Construction of the HIR expression plasmid a *Sal*I restriction fragment containing the entire HIR protein coding sequence[2] was introduced into the *Sal*I site in the polylinker region of the plasmid pUC12 DNA was amplified and molecules containing the 5′ end of the HIR coding sequence next to the pUC12 *Xba*I site were used to generate a *Xba*I-*Dra*I restriction fragment containing the complete HIR protein coding sequence This fragment was introduced into the mammalian expression vector (Fig 1*a*) at an *Xba*I and a filled-in *Bam*HI site in the orientation *Xba*I/*Xba*I-*Dra*I/*Bam*HI Construction of the IER expression plasmid a 931 bp *Bam*HI-*Aat*II restriction fragment from the HIR expression plasmid and a 1,150 bp *Apa*I-*Sst*I restriction fragment of the human EGF receptor sequence[1] were joined at their respective *Aat*II and *Apa*I sites with the complementary synthetic oligonucleotides 5′-CCCGTCAAATATCGCCACTGGGATGG-TGGGGGCC-3′ and 5′-CCCACCATCCCAGTGGCGATATTT-GACGGGACGT-3′ The ligated fragments were inserted into a *Sst*I/*Bam*HI digested pUC12 vector, giving a combination of the sequences coding for the extracellular domain of the insulin receptor with the transmembrane and cytoplasmic domains of the EGF receptor A 965 bp *Bam*HI-*Apa*I fragment of this plasmid, containing the IER junction, was then ligated to a 3 1 kb *Pvu*I-*Bam*HI fragment of the HIR expression plasmid, which contained the rest of the HIR coding sequence and parts of the expression vector, and to two fragments (810 bp *Apa*I-*Bgl*II and 5 3 kb *Bgl*II-*Pvu*I) coding for the transmembrane and cytoplasmic domains of the human EGF receptor, with the latter also containing sequences of the expression vector

---

* Biotechnology Research Center, Meloy Labs, 4 Research Court, Rockville, Maryland 20850, USA

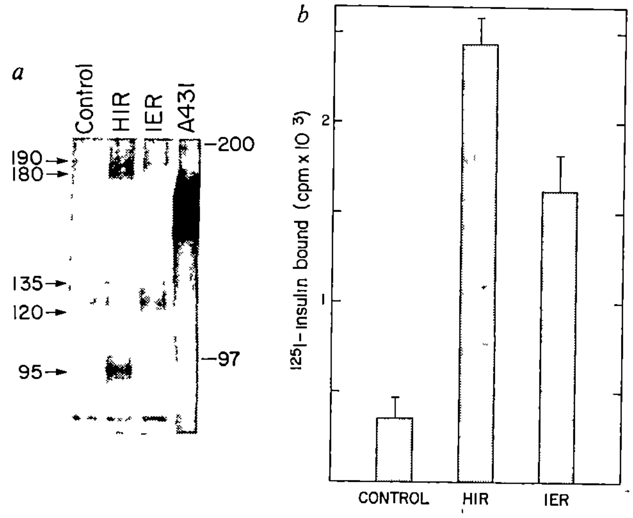


Fig. 2 a, ^{35}S-methionine labelled cDNA expression products. b, ^{125}I-insulin binding by chimaeric receptors. a, COS-7 cells were transfected with the constructs indicated (10 μg supercoiled DNA) or with a control expression vector lacking the receptor cDNA insert[21]. A431 epidermoid carcinoma cells were used for EGF receptor control experiments. COS-7 cells 44 h after transfection were washed twice with methionine-free medium prepared from a MEM select amine kit (Gibco) and incubated for 9 h with 200 μCi ^{35}S-methionine at 800 Ci mmol^{-1} per 8 cm dish. Proteins were immunoprecipitated with monoclonal antibodies to human insulin receptor[22] (MAb CII25.3) or human EGF receptor[23] (MAb RI), as described[24], and were analysed on 7.5% reducing SDS polyacrylamide gels. Sizes of polypeptide markers are indicated (myosin H-chain, 200 K, phosphorylase B, 97.4 K). Arrows, immunoprecipitated expression products. b, Transfected COS-7 cells were washed twice with phosphate-buffered saline (PBS) after 53 h and incubated in 1 ml per 2.2 cm well of serum-free cell culture medium containing 0.2% bovine serum albumin (Sigma), bacitracin (0.5 mg ml^{-1}) (Sigma) and ^{125}I-labelled insulin (0.5 μCi per well) at 93 μCi μg^{-1} (provided by Dr Carl Grunfeld, UCSF) for 2 h at 21 °C. Cells were washed 3 times with PBS at 4 °C, lysed in 0.5 ml 0.1% SDS, 0.1 M NaOH for 30 min at 37 °C, and the total radioactivity was measured. Average of three determinations shown.

brane sequences of the EGF receptor fused to the insulin receptor β subunit extracellular portion (IβE), and, if cleavage of IER proreceptor sequences is inefficient, a 190 K polypeptide. Each of these products was detected in cells expressing the IER chimaera, the IβE fusion glycoprotein subunit migrates as a diffuse band just below the 135 K α-subunit polypeptide (Fig. 2a, see also Fig. 3).

Binding experiments with ^{125}I-labelled insulin (Fig. 2b) indicated that the products of the HIR and IER constructs were expressed in functional form on the surface of COS cells. Insulin binding was highest for the HIR construct and the amount of binding detected was proportional to the levels of the glycopolypeptides present in the transfected cells (Fig. 2a).

To test their intrinsic kinase activity, the expressed polypeptides were immunoprecipitated from lysates of transfected COS-7 cells using antibodies against the EGF receptor (Fig. 3a) or the insulin receptor (Fig. 3b,c), and [γ-^{32}P]ATP was added. Insulin-stimulated autophosphorylation of the human insulin receptor β subunit was clearly detected above the endogenous COS-7 control, at substrate concentrations (100 μM ATP) required by the insulin receptor kinase[15,16] that result in reduced specific activity of the radioactive tracer ([γ-^{32}P]ATP) and thus a weak signal (HIR, Fig. 3b). In contrast, only subpicomolar (0.1 pM) concentrations of ATP were needed to detect EGF-stimulated EGF-receptor kinase activity[17] (Fig. 3a,c), conditions under which the insulin receptor kinase is inactive.

As shown in Fig. 3a and c, the IER chimaeric receptor displays insulin-stimulated autophosphorylation activity. Insulin stimu-

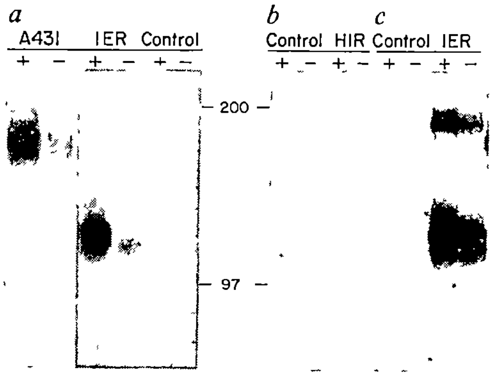


Fig. 3 Autophosphorylation of normal and chimaeric receptors. Immunoprecipitation from a, A431 cells using human EGF receptor mouse monoclonal antibody RI[23], or from COS-7 cells transfected with the IER construct or with a control expression vector lacking a receptor cDNA insert, using human EGF receptor rabbit polyclonal antibody RK-2 (ref. 24), b, HIR-transfected or control cells using human insulin receptor mouse monoclonal antibody CI125.3, c, IER-transfected or control cells using CI125.3. Detergent cell lysates were stimulated with insulin (+) (EGF for A431 cells) or mock-stimulated with PBS (−).

Methods Insulin or EGF stimulation of autophosphorylation was induced by incubating the cell lysates before immunoprecipitation with either 200 nM insulin (Sigma) for 1 h or 1 μM EGF for 15 min at 4 °C. Antibody RI, RK-2 or CI125.3 prebound for 30 min to protein A-Sepharose was added, and incubated for 15 min at 4 °C. The washed immunoprecipitates, in a volume of 30 μl, were adjusted to 5 mM MnCl$_2$, and 15 μCi of [γ-^{32}P] ATP was added for 1 min at 4 °C. The final ATP concentration was 0.1 pM (a and c) or 100 μM (b). The reaction was stopped by adding 20 μl of 3 times concentrated SDS sample buffer, and samples were analysed on 5%/7% SDS polyacrylamide gels[25].

lation of chimaeric protein autophosphorylation was observed in immunoprecipitates obtained using antibodies to either the insulin receptor extracellular domain (Fig. 3c) or the EGF receptor intracellular domain (Fig. 3a). This experiment rules out the possibility that small amounts of insulin-stimulated, endogenous COS cell insulin receptor catalysed phosphorylation of the chimaeric receptors in the immunoprecipitates. The chimaeric IER kinase was active at low ATP concentrations (0.1 pM) consistent with the enzymatic characteristics of the EGF receptor kinase domain. Interestingly, phosphorylation of the bands representing the IβE fusion glycoprotein subunit (120 K) and the uncleaved chimaeric IER proreceptor (190 K) was also induced by insulin. Like the native EGF receptor[17], a maximal ligand induction was observed after 30 seconds at 4 °C (not shown). Thus the EGF receptor tyrosine kinase retains its original enzymatic characteristics when controlled by the insulin binding domain of a chimaeric receptor molecule.

We have described here the expression and functional characterization of a chimaeric receptor molecule composed of the extracellular insulin receptor sequences linked to the EGF receptor transmembrane and cytoplasmic domains. We have shown that binding of insulin to this IER chimaeric receptor molecule induces tyrosine kinase activity in the EGF receptor-derived intracellular domain. The observed tyrosine kinase activity is characteristic of the EGF receptor in that it was measurable at subpicomolar ATP concentrations, which do not activate the insulin receptor phosphotransferase[15,16]. Therefore, the IER chimaeric receptor precursor must have been correctly processed and folded into a conformation permitting communication between two entirely heterologous receptor domains, at the same time preserving the kinetic properties of the EGF receptor tyrosine kinase.

These results directly demonstrate that insulin and EGF receptors use a common transmembrane signal transfer mechanism

This finding was unexpected as the primary structure and, indeed, the general architecture of the extracellular domains of these two proteins show little similarity The only extracellular structures that are homologous between these receptors are regions with high cysteine content and conserved cysteine arrangement[2] that occur twice in the EGF receptor and once per insulin receptor α subunit

How insulin and EGF receptors transduce a signal across the plasma membrane is not fully understood It has been suggested for both receptors that the activation of the protein kinase is likely to be controlled by the state of receptor aggregation, which may be a consequence of ligand binding What alternative mechanisms of transmembrane signal transfer exist for this receptor type? Domain communication through a transmembrane α helix is difficult to envisage, as is interaction of the homologous, cysteine-rich units with a common, heterologous membrane component needed for signal transfer We favour the notion that a ligand-induced conformational change in the receptor extracellular domain triggers receptor dimerization resulting in activation of intracellular enzymatic activities

Whereas the EGF receptor may have to self-associate to induce kinase activity on the cytoplasmic side of the membrane, insulin and IGF-1 receptors[18], both disulphide-linked heterotetramers, may be structured in this way to avoid the need for aggregation and predispose these receptors to respond to ligand The ability of the chimaeric receptors to transduce a signal suggests that the structural differences between EGF and insulin receptors are variations on a common mechanism for signal transfer These results open new avenues of investigation into the mechanisms of transmembrane signalling and receptor function

We thank Drs Ora Rosen and Michael D Waterfield for the supply of monoclonal antibodies, Dr Carl Grunfeld for a gift of ^{125}I-labelled insulin, and M Vasser and P Jhurani for synthesis of oligonucleotides We are indebted to Dr Suzanne Pfeffer for discussions and to Jeanne Arch for help in preparing the manuscript

Received 1 May, accepted 8 September 1986

1 Ullrich, A *et al. Nature* **309**, 418-425 (1984)
2 Ullrich, A *et al. Nature* **313**, 756-761 (1985)
3 Ebina, Y *et al. Cell* **40**, 747-758 (1985)
4 Kahn, C R *A Rev Med* **36**, 429-451 (1985)
5 Herschman, H R in *Control of Animal Cell Proliferation* (eds Boynton, A L & Leffert, H L) 169-199 (Academic, New York 1985)
6 Cohen, S, Fava, R. A & Sawyer, S T *Proc. natn. Acad Sci. USA* **79**, 6237-6241 (1982)
7 Kasuga, M, Hedo, J A, Yamada, K M & Kahn, C R. *J biol Chem.* **257**, 10392-10399 (1982)
8 Hedo, J A, Kasuga, M, Van Obberghen, E, Roth, J & Kahn, C R. *Proc. natn Acad. Sci. USA* **78**, 4791-4795 (1981)
9 Massague, J, Pilch, P F & Czech, M P *J biol. Chem.* **256**, 3182-3190 (1981)
10 Van Obberghen, E *et al. Proc. natn. Acad. Sci. USA* **78**, 1052-1056 (1981)
11 Fujita-Yamaguchi, Y *J biol Chem.* **259**, 1206-1211 (1984)
12 Kasuga, M, Kahn, C R., Hedo, J A, Van Obberghen, E & Yamada, K. M *Proc. natn. Acad Sci. USA* **78**, 6917-6921 (1981)
13 Reed, B C, Ronnett, G V & Lane, D *Proc. natn Acad Sci. USA* **78**, 2908-2912 (1981)
14 Deutsch, P J, Wan, C F, Rosen, O M & Rubin, C S *Proc. natn. Acad Sci. USA* **80**, 133-136 (1983)
15 White, M F, Haring, H-U, Kasuga, M & Kahn, C R. *J biol. Chem.* **259**, 255-264 (1984)
16 Petruzzelli, L, Herrera, R. & Rosen, O M *Proc. natn Acad Sci. USA* **81**, 3327-3331 (1984)
17 Emeux, C, Cohen, S & Garbers, D *J biol Chem* **258**, 4137-4142 (1983)
18 Massague, J & Czech, M P *J biol. Chem.* **257**, 5038-5045 (1982)
19 Simonsen, C C & Levinson, A D *Proc. natn Acad Sci. USA* **80**, 2495-2499 (1983)
20 Lusky, M & Botchan, M *Nature* **293**, 79-81 (1981)
21 Graham, F L & van der Eb, A J *Virology* **52**, 456-467 (1973)
22 Ganguly, S, Petruzzelli, L M, Herrera, R., Stadtmauer, L & Rosen, O M *Curr Top cell. Regul.* **27**, 83-94 (1985)
23 Waterfield, M D *et al. J cell Biochem.* **20**, 149-161 (1982)
24 Kris, R M *et al. Cell* **40**, 619-625 (1985)
25 Laemmli, U K *Nature* **227**, 680-685 (1970)

Cloning of complementary DNA encoding T-cell replacing factor and identity with B-cell growth factor II

Tatsuo Kinashi, Nobuyuki Harada*, Eva Severinson†, Toshizumi Tanabe, Paschalis Sideras†, Mikio Konishi, Chihiro Azuma, Akira Tominaga*, Susanne Bergstedt-Lindqvist†, Masazumi Takahashi, Fumihiko Matsuda, Yoshio Yaoita, Kiyoshi Takatsu* & Tasuku Honjo

Department of Medical Chemistry, Kyoto University Faculty of Medicine, Yoshida, Sakyo-ku, Kyoto 606, Japan
* Department of Biology, Institute for Medical Immunology, Kumamoto University Medical School, Kumamoto 860, Japan
† Department of Immunology, Stockholm University, S-106 91 Stockholm, Sweden

Proliferation and maturation of antigen-stimulated B cells are regulated by several soluble factors derived from macrophages and T cells[1,2]. These soluble factors are functionally divided into two groups: B-cell growth factor (BCGF), thought to be involved in B-cell proliferation; and B-cell differentiation factor (BCDF), responsible for maturation of activated B cells into immunoglobulin-secreting cells[3–6]. This classification needs to be re-examined in the light of the recent cloning of complementary DNA encoding IgG$_1$ induction factor (interleukin-4, IL-4) from the 2.19 mouse T-cell line[7]. Recombinant IL-4 has BCGF and BCDF activities and affects B cells, T cells and mast cells (refs 7, 8; our unpublished data). Another well-characterized B-cell factor is T-cell replacing factor (TRF)[9–12], which, when secreted by the murine T-cell hybridoma B151K12, is defined by two activities[10–12]: induction of IgM secretion by BCL$_1$ leukaemic B-cell line; and induction of secondary anti-dinitrophenol (DNP) immunoglobulin G (IgG) synthesis *in vitro* by DNP-primed B cells Although TRF from B151K12 was classified as BCDF, purified TRF has BCGF-II activity[13]. To elucidate the molecular properties of TRF we isolated cDNA encoding TRF from the 2.19 T-cell line and report here the structure and multiple activities of this lymphokine.

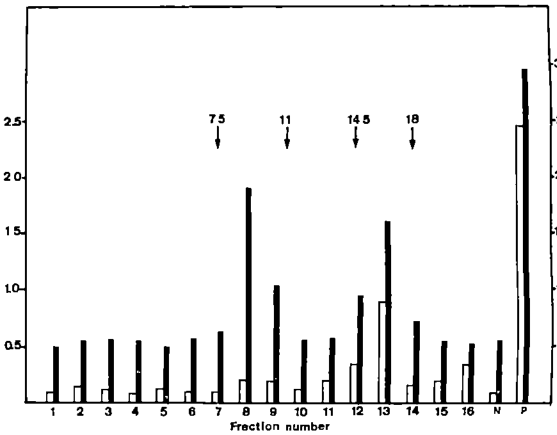


Fig. 1 Effects on BCL$_1$ cells of translational products directed by size-fractionated mRNA of 2 19 T-cell line Poly(A)$^+$ RNA was obtained from the Concanavalin A-activated 2 19 T-cell line and fractionated by centrifugation in a 5–22% sucrose gradient at 36,000 rpm for 15 h[7] An aliquot of each of the 16 fractions was injected into *Xenopus* oocytes Incubation media were collected after incubation for 36 h at 20 °C and added to the assay medium at the concentration of 1% Right-hand scale, ^{3}H-thymidine incorporation (closed bars, c p m per culture × 10^{-3}) Left-hand scale, anti-IgM plaque-formation response (open bars, plaque-forming cells) per culture × 10^{-3} of BCL$_1$ cells, measured as described in Table 1 P, positive control where culture supernatants of B151 hybridoma were added to the medium at the concentration of 50%, N, negative control where phosphate buffer was injected into *Xenopus* oocytes instead of mRNA Arrows, sizes of mRNA (in Svedberg units)

Table 1 TRF and BCGF-II activities on BCL_1 of supernatants of cell culture and mRNA-injected oocytes

Cells of supernatants or templates of mRNA injected into oocytes	Supernatants added (%)	IgM plaque-forming cells per culture	Thymidine incorporation (c p m per culture)
B151 hybridoma	25	1,834	7,069
	10	903	4,222
2 19 T-cell line	1 0	3,134	13,726
	0 25	1,716	6,847
HeLa cell	10	279	705
None	0	211	1,289
rIL-4 (pSP6KmIL4-374)	0 25	32	17,396
B151 TRF	25	1,159	21,242
None	0	60	4,822
*Sac*I-digested library	5	500	8,396
*Sal*I-digested library	5	799	14,342
Phosphate buffer	5	563	6,732
cDNA clone 16	5	199	1,728
18	5	233	1,969
23	5	3,132	17,970
27	5	366	2,001
42	5	215	1,650
51	5	201	1,862
53	5	180	1,883
none	5	150	1,654
*Sal*I-digested clone-23	1	3,229	10,824
	0 25	1,662	6,615
	0 06	772	3,037
	0 015	232	1,625
None	0	211	1,289
*Acc*I-digested clone-23	10	2,343	18,772
*Sal*I-digested clone-23	10	2,277	23,386
None	0	208	4,904

Supernatants of cells or oocytes injected with mRNA were added to the culture medium of BCL_1 cells at indicated concentrations BCL_1 cells ($1 5\times10^5$) of an *in vivo* line were cultured in 0 2 ml RITC medium (serum-free, 5 mg ml^{-1} bovine serum albumin, 50 μM 2-mercaptoethanol and antibiotics) in addition to testing supernatants After a 2-day culture, the protein A plaque assay[19] was used to determine the numbers of IgM-secreting cells BCL_1 proliferation was determined by ^3H-thymidine uptake for the last 6 h of a 2-day culture Culture medium of various cell lines were added to the assay medium at indicated concentrations rIL-4, Supernatants of *Xenopus* oocytes injected with mRNA derived from pSP6KmIL4-374 as described previously[7] B151 TRF prepared by partial purification from supernatants of B151 cells[13] The total cDNA library consisted of 5×10^4 clones and the library DNAs digested with *Sac*I or *Sal*I The detailed methods of mRNA synthesis by SP6 RNA polymerase and its injection to *Xenopus* oocytes were described previously[7] Individual cDNA clones with 1-kb longer inserts containing *Sac*I sites were digested with *Sal*I and their mRNAs tested Incubation medium of oocytes injected with distilled water was used as a negative control The construct pSP6K-mTRF23 was digested either by *Sal*I or by *Acc*I and used as template for *in vitro* mRNA synthesis All values are means in the triplicate experiments and standard errors of triplicate experiments ranged from 1 01 to 1 12

Supernatants of the 2 19 T-cell line[14] contain the TRF activity (Table 1) and, like the B151 supernatants, have the BCGF-II activity as assayed by induction of proliferation of BCL_1 cells $Poly(A)^+RNA$ of 2 19 T cells was fractionated by sucrose density gradient centrifugation and an aliquot of each fraction was injected into *Xenopus* oocytes to test the presence of messenger RNA encoding BCGF-II and TRF by measuring proliferation stimulation and IgM plaque formation of BCL_1 cells, respectively BCL_1 proliferation occurs in translation products directed by 9 and 16S mRNAs (Fig 1) In contrast, IgM plaque-forming activity is found only in products of 16S mRNA The 9S mRNA corresponds to the size of mRNA for IgG_1 induction factor (IL-4)[7] Recombinant IL-4 produced by *Xenopus* oocytes shows BCL_1 proliferation activity but not IgM plaque-forming activity (Table 1) Furthermore, translation products directed by 15–18S mRNA of B151K12 cells exert TRF activity[15] We therefore conclude that 16S mRNA of 2 19 T cells encodes another factor that has TRF activity The same fraction of mRNA also encodes BCGF-II defined by the growth stimulation of BCL_1 The results also indicate that the BCL_1 proliferation assay does not discriminate between BSF-I (IL-4) and BCGF-II

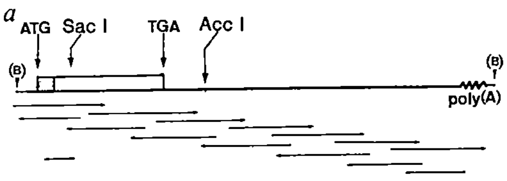
To clone cDNA encoding TRF we used the pSP6K cDNA library of the total $poly(A)^+$ RNA of 2 19 T cells, the construction of which was described previously[7] A DNA mixture of 5×10^4 independent cDNA clones was cleaved with *Sac*I or *Sal*I to linearize plasmid DNA and mRNAs were synthesized using SP6 RNA polymerase The synthesized RNAs were injected into

Table 2 Stimulation of anti-DNP IgG plaque formation by rTRF

Pretreatment with anti-Thy 1 2	Antigen	TRF in culture (%)	Anti-DNP IgG plaque-forming cells per culture
–	None	None	127
–	DNP-KLH	None	1,792
+	None	None	127
+	DNP-KLH	None	199
+	DNP-OA	None	92
+	DNP-OA	rTRF 1	695
+	DNP-OA	rTRF 0 5	884
+	DNP-OA	rTRF 0 25	555
+	DNP-OA	rTRF 0 125	382
+	DNP-OA	B151 10	720
+	DNP-OA	B151 5	636
+	DNP-OA	B151 2 5	485
+	DNP-OA	B151 0 25	325

BALB/c mice were immunized with DNP-KLH and spleen cells isolated 6–8 weeks after the immunization and treated twice with anti-Thy 1 2 and complements The cells (6×10^5) were cultured in 0 2 ml culture media (RPMI 1640) containing 10% fetal calf serum (FCS) together with 10 ng DNP-ovalbumin (DNP-OA) and TRF for 5 days Anti-DNP IgG plaque-forming cells were measured as described[11] B151 indicates partially purified TRF from supernatants of B151 cells as described[13]

a



c

```
           41                                74
mTRF       ALLTSNETMRLPVPTHKN-HQLCIGEIFQGL-DI-LK
           *  ,   * **   *   *      . **   *   **
mIL-3      ASISGRDTHRLT-RTL-NCSSI-VKEIIGALPEPELK
           27                                60

           87                   108
mTRF       FQNLSLIKKYIDRQKEKCGEER
           *    *  *    * *   * *
mIL-3      FRRVNL-SKFVESQGEVDPEDR
           73                   93
```

```
           67         78
mTRF       FQGL-DILK-NQTV
             *   ** **
mIFN-γ     YLRLFEVLKDNQAI
           74           87

           114            128
mTRF       FLDYLQEFLG-VMSTE
           * * *  **      *
mGM-CSF    FIDSLKTFLTDI-PFE
           97             111
```

b

```
                                          1                             10                                                    20
                                          Met Arg Arg Met Leu Leu His Leu Ser Val Leu Thr Leu Ser Cys Val Trp Ala Thr Ala Met Glu
CGCTCTTCCTTTGCTGAAGGCCAGCGCTGAAGACTTCAGAGTC ATG AGA AGG ATG CTT CTG CAC TTG AGT GTT CTG ACT CTC AGC TGT GTC TGG GCC ACT GCC ATG GAG
1                                                                                                                           100

                     30                                  40                              50
Ile Pro Met Ser Thr Val Val Lys Glu Thr Leu Thr Gln Leu Ser Ala His Arg Ala Leu Leu Thr Ser Asn Glu Thr Met Arg Leu Pro Val Pro Thr
ATT CCC ATG AGC ACA GTG GTG AAA GAG ACC TTG ACA CAG CTG TCC GCT CAC CGA GCT CTG TTG ACA AGC AAT GAG ACG ATG AGG CTT CCT GTC CCT ACT
                                                                                                                            200

                60                                  70                                  80
His Lys Asn His Gln Leu Cys Ile Gly Glu Ile Phe Gln Gly Leu Asp Ile Leu Lys Asn Gln Thr Val Arg Gly Gly Thr Val Glu Met Leu Phe Gln
CAT AAA AAT CAC CAG CTA TGC ATT GGA GAA ATC TTT CAG GGG CTA GAC ATA CTG AAG AAT CAA ACT GTC CGT GGG GGT ACT GTG GAA ATG CTA TTC CAA
                                                                                                                            300

 90                                 100                                 110                                 120
Asn Leu Ser Leu Ile Lys Lys Tyr Ile Asp Arg Gln Lys Glu lys Cys Gly Glu Glu Arg Arg Arg Thr Arg Gln Phe Leu Asp Tyr Leu Gln Glu Phe
AAC CTG TCA TTA ATA AAG AAA TAC ATT GAC CGC CAA AAA GAG AAG TGT GGC GAG GAG AGA CGG AGG ACG AGG CAG TTC CTG GAT TAC CTG CAA GAG TTC
                                                                                                                             400

                        130
Leu Gly Val Met Ser Thr Glu Trp Ala Met Glu Gly END
CTT GGT GTG ATG AGT ACA GAG TGG GCA ATG GAA GGC TGA GGCTGAGCTGCTCCATGGTGACAGGACTTCACAATTTAAGTTAAATTGTCAACAGATGCAAAAACCCCACAAAACTGTG
                                                                                                               500

CAAATGCAAGGGATACCATATGCTGTTTCCATTTATATTTATGTCCTGTAGTCAGTTAAACCTATCTATGTCCATATATGCAAAGTGTTTAACCTTTTTGTATACCCATAAAAGAAATTCCTGTAGCGCAC
                                                                              600

CCTGGCCTCAAACTGGTAATGTAGCCAAGGATAACCTTGAATTTCTGATCCTCCTGCCTCCTCTTCCTGAAGGCTGAGGTTACAGACATGCACCATTGCCACTAGTTCATGAAGTGCTGCAGATGGAACCC
                                              700

AAGGCTTTGTGCATGTTACCAACTGAGTTATACTCCCTCCCCCTCATCCTCTTCGTTGCATCAGGGTCTCAAGTATTCCAGGCTGACTTGAACTCAGTGTGTAGCCAAGGGTGACCCTGAACTCTTGGTC
                800                                                                                                        900
CAGATGGACGCAGGAGGATCACATACCCAACCTTAGCATCCTTTCTCCTAGCCCCTTTAGATAGATGATACTTAATGACTCTCTTGCTGAGGGATGCCACACCGGGGCTTCCTGCTCCTATCTAACTTCAA
                                                                                      1000

TTTAATACCCACTAGTCAATCTCTCCTCAACTCCCTGCTACTCTCCCCAAACTCTAGTAAGCCCACTTCTATTTCTTGGGGACAGAGAAGGTTGACTTTTCTTATGTCCTATGTATGAATCAGACTGTGCC
                                                      1100

ATGACTGTGCCCTCTGTGCCTGGAGCAGCTGGATTTTGGAAAAGAAAAGGGACATCTCCTTGCAGTGTGAATGAGAGCCACCCACATGCTGGGCCTTACTTCTCCGTGTAACTGAACTTAAGAAGCAAAGTA
                       1200                                                                                             1300
AATACCACAACCTTACTACCCCATGCCAACAGAAAGCATAAAATGGTTGGGATGTTATTCACGTATCAGGGTCACTGGAGAAGCCTCCCCCAGTTTACTCCAGGAAAAACAGATGTATGCTTTTATTTAAT
                                                                                              1400

TCTGTAAGATGTTCATATTATTTATGATGGATTCAGTAAGTTAATATTTATTACACGTATATAATATTCTAATAAAGCAGAAGGGACAACTC
                                                                                             (A) ∼90
                                                              1500
```

Fig 2 *a*, Restriction map of the insert of pSP6K-mTRF23 and its sequence strategy The insert of pSP6K-mTRF23 was isolated by *Bam*HI digestion and subcloned in pUC18 Clones that contained series of unidirectional deletions were produced by digestion with exonucleases III and VII from either end[16] The nucleotide sequence of the entire cDNA was determined by the dideoxy method[17] Top line, insert DNA, dotted and open rectangles, DNA sequences encoding the signal peptide and coding region, respectively Horizontal arrows, directions and ranges of sequences read Double arrow, 100 bp B, *Bam*HI site *b*, Nucleotide and predicted amino-acid sequences of TRF Possible N-glycosylation sites are underlined Numbers above and below sequences, amino-acid and nucleotide positions, respectively *c*, Regions of identity between TRF and other lymphokines Similar amino-acid sequences of known proteins were searched by computer program Sequences (>14 residues) with 45% similarity (including related amino acids at the corresponding region) were chosen Identical and related amino acids are shown by asterisks and dots, respectively Sequences of murine GM-CSF[21], murine interferon-γ (IFN-γ)[22] and murine IL-3[23,24] were taken from published details

Xenopus oocytes and the translation products secreted into the incubation medium collected and assayed for proliferation and plaque-formation of BCL$_1$ cells. Translation products derived from *Sal*I-digested templates have both of these activities whereas those from *Sac*I-digested templates have neither, indicating that cDNAs encoding BCGFII and TRF contain *Sac*I sites in the structural sequences (Table 1)

The pSP6K cDNA library of 2 19 T-cell mRNA was divided into 18 pools, each consisting of 3,000 clones We assayed each pool and found only one that was positive both in proliferation and in IgM plaque-forming activity of BCL$_1$ cells After two more subpooling steps, we identified a positive pool of 60 clones and examined inserts of the individuals to select clones whose inserts were >1 kilobase (kb) and contained the *Sac*I site but not the *Sal*I site Seven out of 60 clones (clones 16, 18, 23, 27, 42, 51 and 53) fulfilled these criteria One of them, clone 23, was positive for BCL$_1$ proliferation and IgM plaque-formation (Table 1) and was termed pSP6K-mTRF23

The nucleotide sequence of the entire cDNA was determined by the combination of the unidirectional deletion method[16] using exonucleases III and VII and the dideoxy method[17] The cDNA insert is 1,533 base pairs (bp) long excluding the poly(A) tail (Fig 2) There are two long open reading frames (ORFs) deduced from the cDNA sequence, the longer frame (133 residues) between bases 44 and 442 and the shorter (62) between bases 544 and 729 The TRF activity occurs in the translation products of pSP6K-mTRF23 linearized by *Acc*I (Table 1) Because the *Acc*I site is located in the middle of the shorter ORF at bases 624-629, the functional significance of this reading frame was excluded The longer ORF encodes a polypeptide of 133 amino acids with three possible *N*-glycosylation sites There are 21 hydrophobic amino acids at the N-terminus of this ORF

Table 3 BCGFII activity of recombinant TRF

Lymphokines	Additions (%)	c.p.m. per culture DXS −	c.p.m. per culture DXS +	Synergy indices
none		520	842	1.6
rTRF	5	2,715	6,044	2.2
rTRF	1.25	1,724	4,268	2.5
rTRF	0.6	1,290	2,813	2.2
rTRF	0.16	919	1,827	2.0
2.19 sup	3.0	7,425	11,214	1.5
2.19 sup	1.0	7,183	10,740	1.5
2.19 sup	0.3	6,424	8,626	1.3

CBA spleen B cells enriched by panning[20] were cultured in duplicates of $2\times10^6\,\mathrm{ml}^{-1}$ in 0.2 ml RPMI1640 containing 50 μM 2-mercaptoethanol and 15% FCS with or without (50 μg ml^{-1}) dextran sulphate (DXS). Supernatants of oocytes injected with pSP6K-mTRF23 mRNA (rTRF) and culture supernatants of 2.19 T cell line (2.19 sup) were added at indicated concentrations. ^{3}H-thymidine was added for the last 24 h of a 72-h culture and the degree of proliferation was assessed by determining incorporation of the radiolabel. Synergy indices were calculated by dividing c.p.m. in DXS-stimulated cultures by c.p.m. in unstimulated culture.

(including two arginine residues). The putative secreted core polypeptide has a relative molecular mass (M_r) of 12,300, which is near that of TRF (M_r 18,000 after glycosylation[12]). The longer ORF contains three cysteine residues (15, 62 and 104) that may be involved in the oligomer formation. We therefore conclude that the longer ORF encodes the TRF molecule.

We searched homologies between the deduced amino-acid sequence of TRF and sequences of known proteins, including all the lymphokines so far identified. No proteins are very similar to TRF except for short segments of murine interleukin-3 (IL-3), murine granulocyte–macrophage colony stimulating factor (GM-CSF) and murine interferon-γ (Fig. 2c). We suggest that TRF is distantly related to these lymphokines.

We tested various biological activities of recombinant TRF (rTRF), the oocyte translation product of pSP6K-mTRF23. First, we tested it for the other activity of TRF, namely, stimulation of the secondary anti-DNP IgG response. Mice were immunized with DNP-keyhole limpet haemocyanin (KLH), and splenic B cells were purified 6–8 weeks later for the assay of the anti-DNP IgG plaque-forming cells. The assay revealed that rTRF stimulates DNP-primed B cells in the presence of antigen (DNP-OA) and augments anti-DNP IgG response (Table 2). The rTRF also induces IgM synthesis in *in vivo*-activated B-cell blasts (data not shown). These properties of rTRF agree with the criteria of TRF derived from B151 (refs 11–13).

During screening, we found that pools containing the stimulating activity of IgM plaque-forming cells of $\mathrm{BCL_1}$ cells always have the stimulation activity of the $\mathrm{BCL_1}$ proliferation. The translation product of the final clone, pSP6K-mTRF23, also shows these two activities (Table 1). Furthermore, rTRF augments ^{3}H-thymidine incorporation into B cells stimulated with dextran sulphate (Table 3). These results indicate that the TRF and BCGFII activities belong to a single molecule. The rTRF has no interleukin-1 (IL-1)-, IL-2-, IL-3- or BSF-I-like activity (data not shown), but it induces augmentation of ^{3}H-thymidine uptake in thymocytes in conjunction with suboptimal doses of Concanavalin A and IL-2. Although Sanderson *et al.*[18] have recently suggested that BCGF-II is identical to eosinophil differentiation factor, we do not yet know whether TRF has this activity. Because of the diverse activities and targets of TRF, this lymphokine is classified as an interleukin. As this is the fifth interleukin whose structure has been determined, we propose that TRF or BCGF-II be called interleukin-5.

We thank Professor T. Oi for advice on the homology search and Ms K. Hirano for preparation of the manuscript. Supported by grants from the Ministry of Education, Science and Culture of Japan, the Swedish National Science Research Council, the Swedish MRC and the Aristotale Onassis benefit foundation.

Received 23 June; accepted 20 August 1986.

1. Dutton, R. W., Folkoff, R. & Hirst, J. A. *Prog. Immun.* **1**, 355–368 (1971).
2. Swain, S. L. *et al. J. exp. Med.* **158**, 822–835 (1983).
3. Howard, M., Nakanishi, K. & Paul, E. W. *Immun. Rev.* **78**, 185–210 (1984).
4. Kishimoto, T. *A. Rev. Immun.* **3**, 133–157 (1985).
5. Isakson, P. C., Pure, E., Vitetta, E. S. & Krammer, P. H. *J. exp. Med.* **155**, 734–748 (1982).
6. Bergstedt-Lindqvist, S., Sideras, P., MacDonald, H. R. & Severinson, E. *Immun. Rev.* **78**, 25–50 (1984).
7. Noma, Y. *et al. Nature* **319**, 640–646 (1986).
8. Lee, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 2061–2065 (1986).
9. Schimpl, A. & Wecker, E. *Nature new Biol.* **237**, 15–17 (1972).
10. Takatsu, K., Tominaga, A. & Hamaoka, T. *J. Immun.* **124**, 2414–2422 (1980).
11. Takatsu, K. *et al. J. Immun.* **125**, 2646–2653 (1980).
12. Takatsu, K. *et al. J. Immun.* **134**, 382–389 (1985).
13. Harada, N. *et al. J. Immun.* **134**, 3944–3951 (1985).
14. Sideras, P., Bergstedt-Lindqvist, S., MacDonald, H. R. & Severinson, E. *Eur. J. Immun.* **15**, 586–593 (1985).
15. Tominaga, A., Matsumoto, M., Takahashi, T., Harada, N. & Takatsu, K. *Microbiol. Immun.* (in the press).
16. Yanisch-Perron, C., Vieira, J. & Messing, J. *Gene* **33**, 103–119 (1985).
17. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5469 (1977).
18. Sanderson, C., O'Garra, A., Warren, J. D. & Klaus, G. B. G. *Proc. natn. Acad. Sci. U.S.A.* **83**, 437–440 (1986).
19. Gronowicz, E., Coutinho, A. & Melchers, F. *Eur. J. Immun.* **6**, 588–590 (1976).
20. Mage, M. G., McHugh, L. L. & Rothstein, T. L. *J. immun. Meth.* **15**, 47–56 (1977).
21. Gough, M. N. *et al. Nature* **309**, 763–767 (1984).
22. Gray, W. P. & Goeddel, V. D. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5842–5846 (1983).
23. Fung, C. M. *et al. Nature* **307**, 233–237 (1984).
24. Yokota, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 1070–1074 (1984).

Complementary DNA for a novel human interleukin (BSF-2) that induces B lymphocytes to produce immunoglobulin

Toshio Hirano, Kiyoshi Yasukawa, Hisashi Harada*, Tetsuya Taga, Yasuo Watanabe, Tadashi Matsuda, Shin-ichiro Kashiwamura, Koichi Nakajima, Koichi Koyama, Akihiro Iwamatsu†, Susumu Tsunasawa†, Fumio Sakiyama†, Hiroshi Matsui‡, Yoshiyuki Takahara‡, Tadatsugu Taniguchi* & Tadamitsu Kishimoto

Division of Immunology and * Division of Molecular Biology,
Institute for Molecular and Cellular Biology, † Division of Protein
Chemistry, Institute for Protein Research, Osaka University,
1-3, Yamada-Oka, Suita, Osaka 565, Japan
‡ Central Research Laboratories, Ajinomoto Co. Inc., Totsuka-ku,
Yokohama 244, Japan

When stimulated with antigen, B cells are influenced by T cells to proliferate and differentiate into antibody-forming cells. Since it was reported[1,2] that soluble factors could replace certain functions of helper T cells in the antibody response, several different kinds of lymphokines and monokines have been reported in B-cell growth and differentiation[3,4]. Among these, human B-cell differentiation factor (BCDF or BSF-2) has been shown to induce the final maturation of B cells into immunoglobulin-secreting cells[5–8]. BSF-2 was purified to homogeneity[9] and its partial NH_2-terminal amino-acid sequence was determined[10]. These studies indicated that BSF-2 is functionally and structurally unlike other known proteins. Here, we report the molecular cloning, structural analysis and functional expression of the cDNA encoding human BSF-2. The primary sequence of BSF-2 deduced from the cDNA reveals that BSF-2 is a novel interleukin consisting of 184 amino acids.

In a previous study[11], we established a human T-cell line (TCL-Na1), which constitutively produced large amounts of BSF-2, using human T-cell leukaemia virus-1 (HTLV-1). BSF-2, with a relative molecular mass (M_r) 21,000 (21K), was purified

a

1	S	N	M	X	E	X	X	K			
2	P	V	P	P	G	E	D	S	K		
3	E	A	L	A	E	N	X	X	X	L	X	K
4	N	L	X	X	I	T	X	P			
5	D	G	X	F	Q	X	G	F	N		
6	X	V	A	A	P						
8	L	Q	A	Q	N	Q	W	L	Q	X	M

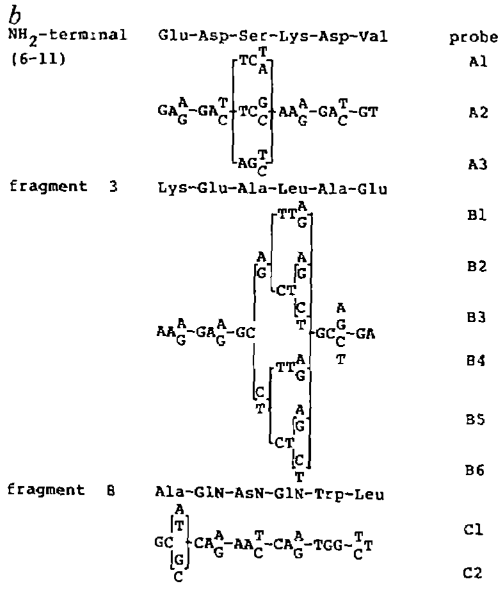


Fig. 1 *a*, Sequences of BSF-2 fragments, *b*, sequences of synthetic oligonucleotide probes

Methods. *a*, BSF 2 (21K) was purified as reported previously[9,10] Approximately 1 nmol of purified BSF-2 (20 μg) was digested with 5 pmol of lysylendopeptidase in 5 mM Tris-HCl buffer (*p*H 9 5) at 37 °C for 6 h The fragments were separated by high performance liquid chromatography (HPLC) using reverse phase column, μBondopack (C_{18} 300×2 1 mm, Waters), with a linear gradient of 0–60% acetonitrile containing 0 1% trifluoroacetic acid The partial amino-acid sequence of each fragment was determined using a gas-phase protein sequencer (Applied Biosystems model 470A) essentially as described[17] The obtained phenylthiohydantoin-amino acid residues were analysed by HPLC using ODS-120T column[18] (C_{18} 250×4 6 mm, Toyo Soda) *b*, Mixtures of 32 (A1–A3, C1 and C2) or 64 (B1–B6) possible oligonucleotides (17 bases) corresponding to the NH_2-terminal (A1–A3), fragment 3 (B1–B6) and fragment 8 (C1, C2) were made by an automated DNA synthesizer (Applied Biosystem Inc) The expression vector pQ was constructed by a series of manipulations of pCEIL-2[19] Essentially, it contains two copies of a simian virus 40 (SV40) DNA segment containing the early promoter and *ori* sequences in opposite orientations This plasmid is unstable unless a DNA (in this case cDNA) segment is inserted between the two SV40 sequences, because it becomes cruciform in *Escherischia coli*[20] Once the cDNA is inserted it can be expressed in monkey COS cells independent of orientation Efficiency of expression was previously confirmed with IL-2 and IFN cDNA (G Yamada and T Tanigushi, unpublished data) Poly (A)$^+$ RNA was prepared from TCL-Na1 cells and cDNA was made using the Amersham cDNA synthesis system (Amersham UK) which is essentially based on the method described by Gubler and Hoffman[21] pQ vector was digested with *Pst*I to remove the insert IL-2 cDNA One part of the cDNA was annealed with pQ vector at the *Pst*I site by the standard G-C tailing method[19] and the resulting recombinant plasmids were transfected into *E coli* MC1061 The other part of the cDNA was methylated with *Eco*RI methylase (New England Biolabs), ligated with *Eco*RI linker (Takara Shuzo Co), digested with *Eco*RI and applied to Bio-Gel A-50 m (Bio-Rad) cDNA fragments longer than 500 base pairs were ligated with *Eco*RI-digested λgt10 vector (Stratagene) and packaged *in vitro* using Gigapack (Stratagene)[22] Both cDNA libraries (30×10^4 clones) were screened with the ^{32}P-labelled 17-base-long oligonucleotides by the *in situ* procedure under the conditions described[23]

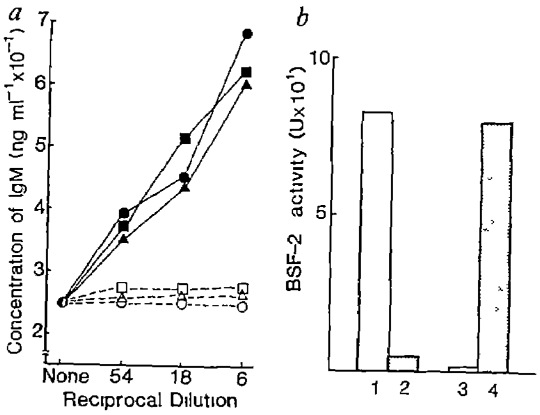


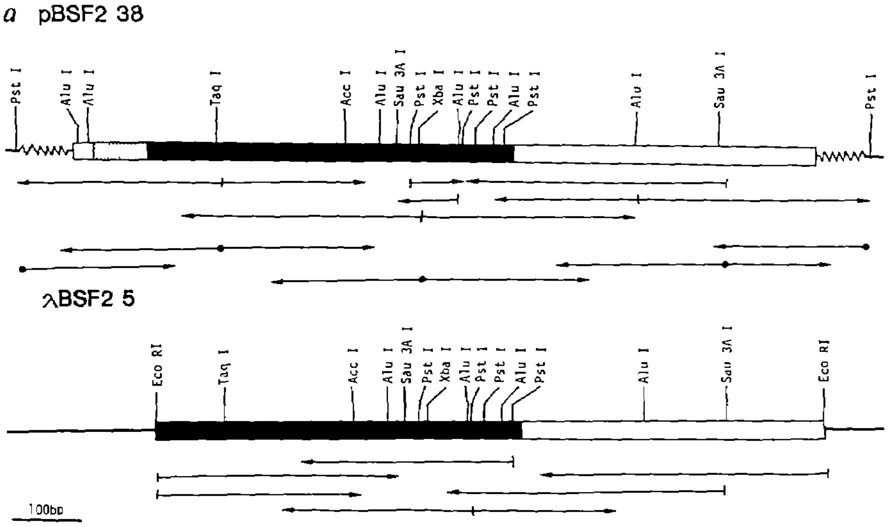
Fig 2 *a*, Expression of BSF-2 cDNA in COS7 cells, *b*, affinity chromatography of COS7 cell-derived BSF-2

Methods Monkey COS7 cells[24] (5×10^5) were seeded onto 60 mm plates one day before transfection Transfections were carried out with 2 μg of plasmid DNA in 1 ml 25 mM Tris-HCl (*p*H 7 4), 137 mM NaCl, 5 mM KCl, 0 6 mM Na_2HPO_4, 0 5 mM $MgCl_2$, 0 7 mM $CaCl_2$ and 500 μg of DEAE-Dextran (Pharmacia) After 1 h incubation at 37 °C, the solution was replaced with Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS) and 150 μM chloroquine, incubated at 37 °C for 3 h and replaced with fresh DMEM containing 10% FCS Seventy-two hours later the culture supernatants were collected and assayed for BSF-2 activity *a*, Ten thousand cells of the Epstein-Barr virus transformed B cell line SKW6-C14 were cultured in different concentrations of either the culture supernatants of COS7 cells transfected with pBSF2 38 DNA (●, ■, ▲) or mock-transfected COS7 cells (○, □, △) Three separate experiments are shown After 3 days the levels of IgM in the culture supernatants were determined by enzyme-linked immunosorbent assay (ELISA) as described[9] *b*, The culture supernatants of COS7 cells transfected with pBSF2 38 DNA (containing 90 units of BSF-2) were applied to specific anti-peptide immunoaffinity gel (2, 4) or to control anti-OVA immunoaffinity gel (1, 3) After washing the gels, the bound protein was eluted with 3M KSCN BSF-2 activities in the effluent (1, 2) and eluate (3, 4) were determined using SKW6-C14 cells One unit per ml of BSF-2 induces 50% of the maximum IgM-production in 1×10^4 SKW6-C14 cells[9] The specific immunoaffinity gel was made using affinity-purified rabbit antibody (antipeptide) against the synthetic NH_2-terminal BSF-2 peptide[10]

to homogeneity from the culture supernatant of TCL-Na1 cells and digested with lysylendopeptidase to give 9 fragments The partial amino acid sequences of 7 fragments were determined (Fig 1*a*) On the basis of the partial amino acid sequence at the NH_2-terminal of BSF-2 reported previously[10], it was shown that fragment 2 consisted of 9 amino acids at the NH_2-terminal and fragment 6 consisted of residues starting with the tenth amino acid from the NH_2-terminal The cDNA libraries made from poly(A)$^+$ RNA of TCL-Na1 cells were probed with 11 groups of synthetic oligonucleotide mixtures (17 bases each) corresponding to fragments 3, 8 and the NH_2-terminal as indicated in Fig 1*b* One cDNA clone, designated pBSF2 38, that specifically hybridized with A2, B2 and C1 probes but no others, was isolated from the cDNA library constructed in pQ vector The libraries were screened again by pBSF2 38 insert cDNA as a probe and ten positive clones were identified Four of these were shown to hybridize with the A2 probe, which corresponds to the NH_2-terminal of BSF-2 All five cloned plasmids contained cDNA inserts consisting of about 1,000–1,100 base pairs To examine whether pBSF2 38 contained the entire coding region of BSF-2, pBSF2 38 DNA was transfected into COS7 cells The culture supernatant of COS7 cells transfected with pBSF2 38 DNA induced IgM production in SKW6-C14 cells and the control culture supernatant of mock-transfected COS7 cells did not (Fig 2*a*) The BSF-2 activity generated by COS7 cells transfected with pBSF2,38 DNA could be adsorbed onto and subsequently eluted from an immunoaffinity gel conjugated with polyclonal antibody against the synthetic NH_2-terminal BSF-2 peptide, but did not bind to the control antiovalbumin (OVA) immunoaffinity gel (Fig 2*b*) Thus pBSF2 38 contains the entire coding region of BSF-2

The complete nucleotide sequence of the insert cDNA of pBSF2 38 was determined by the dideoxy method[12] and by chemical cleavage[13] The restriction endonuclease cleavage map of the cDNA insert and the sequencing strategy are shown in Fig 3*a* The insert cDNA of pBSF2 38 contains a single large open reading frame (Fig 3*b*) In this frame, the initiator ATG is followed by 211 codons before the termination triplet TAG We confirmed the nucleotide sequence of BSF-2 cDNA using one of the clones (λBSF2 5) identified by probing the cDNA library constructed in λgt 10 vector with pBSF2 38 insert DNA (Fig 3*a*) Although λBSF2 5 insert cDNA was found to lack 126 nucleotides at the 5′ terminal region compared with pBSF2 38 insert cDNA, the nucleotide sequence of the other region was in complete agreement with that of pBSF2 38 cDNA insert (data not shown) Because we know from protein sequence data that the sequence of the NH_2-terminal region of mature BSF-2 is Pro-Val-Pro-Pro , we conclude that BSF-2 is made as a precursor consisting of 212 amino acids and processed into a mature form by a cleavage between the Ala (−1) and Pro (+1) residues Hydrophobic amino acids are abundant in residues from −28 to −1 and this region seems to be a typical signal peptide that is required for BSF-2 secretion Although we do not know at present whether the carboxy-terminal processing occurs in BSF-2, we conclude that mature BSF-2 probably consists of 184 amino acids As underlined in Fig 3*b*, the amino-acid sequence of all the fragments that originated from purified BSF-2 was in complete agreement with the amino-acid sequence deduced from the cDNA The calculated molecular mass of the BSF-2 polypeptide was 20,781 5 This is comparable with the apparent molecular mass of purified BSF-2 (21K) as determined by SDS-polyacrylamide gel electrophoresis (PAGE)[9,10] Two potential *N*-glycosylation sites (Asn-X-Ser and Asn-X-Thr, at positions 45–47 and 144–146, respectively) were identified, but we have not yet determined whether BSF-2 is a glycoprotein

The protein sequence of BSF-2 was compared with all proteins in the National Biomedical Research Foundation database and Genetic Sequence Data Bank (Bolt Beranek and Newman, Inc, Los Alamos National Laboratory), no strong similarity with any protein was revealed We also searched our personal database, which contains human IL-1α, IL-1β, IL-2, γ-interferon (γ-IFN), colony stimulating factor-1 (CSF-1), granulocyte macrophate (GM)-CSF, granulocyte (G)-CSF and B-cell stimulatory factor-1 (BSF-1), and murine IL-1, IL-2, γ-IFN, GM-CSF, IL-3 and



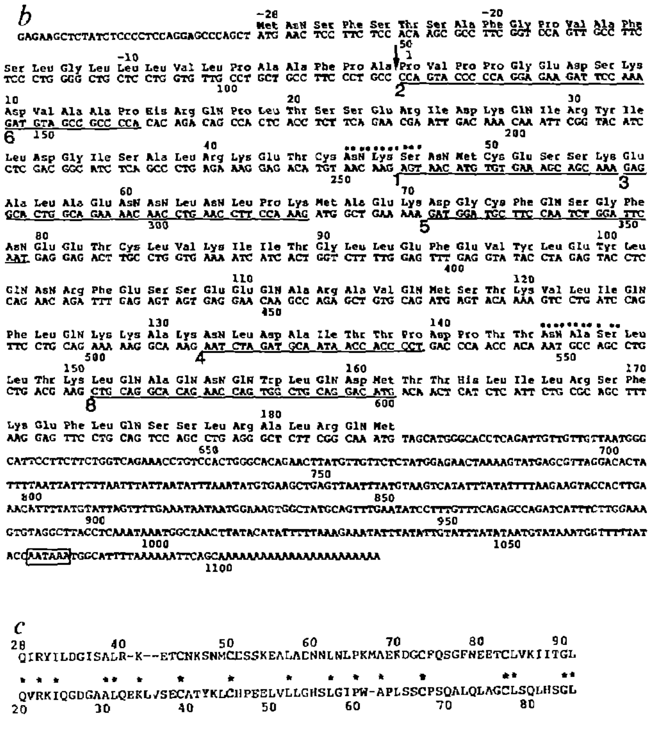

b
c
28 40 50 60 70 80 90
QIRYILDGISALR-K--ETCNKSNMCESSKEALAENNLNLPKMAEKDGCFQSGFNEETCLVKIITGL
QVRKIQGDGAALQEKLVSECATYKLCHPEELVLLGHSLGIPW-APLSSCPSQALQLAGCLSQLHSGL
20 30 40 50 60 70 80

Fig 3 *a*, Restriction endonuclease cleavage map and sequence strategy for the cDNA insert of pBSF2 38 and λBSF2 5 *b*, Nucleotide sequence and deduced amino-acid sequence of the plasmid pBSF2 38 cDNA insert *c*, Homology of human BSF2 and G-CSF *a*, Arrows with vertical marks indicate the direction and extent of sequencing by the dideoxy chain termination method Arrows with dots, regions of DNA sequenced by chemical cleavage Scale, 100 bp *b*, Underlined amino acids, the region matching the sequence of the lysylendopeptidase-digested peptides 1–6 and 8 (see Fig 1*a*) All derived peptide sequences are preceded by a lysine, consistent with lysylendopeptidase digestion Asterisks, potential *N*-glycosylation sites (Asn-X-Ser/Thr), box, presumed poly(A) addition signal sequence Numbers above and below the sequence show positions of amino acids and nucleotides, respectively Amino acids are numbered starting at Pro (1) of the mature BSF-2 protein sequence *c*, Homology search was performed using the search program of Wilbur and Lipmann[25] The parameters used in the database search were κ-tuple size = 2, window size = 20, weight for a deletion = 8 and value to be added to each element of the similarity matrix = 0 Under these conditions, only human G-CSF showed significant similarity to BSF-2 (homology score = −129) Asterisks, homologous amino acids, shaded boxes, cysteine Sequence of G-CSF was taken from Nagata *et al*[14]

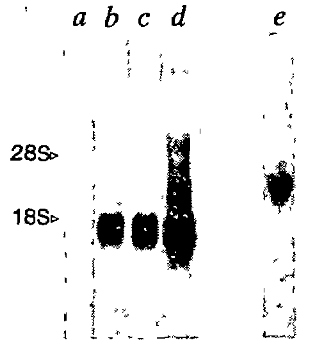


Fig. 4 Analysis of BSF-2 mRNA. Poly(A)$^+$ RNA was prepared from various kinds of cells. RNA (10 µg) was denatured and subjected to blotting analysis[26]. *a*, Poly(A)$^+$ RNA from non-stimulated tonsillar mononuclear cells, *b*, poly(A)$^+$ RNA from tonsillar mononuclear cells stimulated with phytohaemagglutinin (0.1%) and tetradecanoyl phorbol 13-acetate (5 ng ml^{-1}) for 40 h, *c*, poly(A)$^+$ RNA from TCL-Na1 cells, *d*, poly(A)$^+$ RNA from T 24 bladder cell carcinoma, *e*, total RNA from cardiac myxoma cells[10].

BSF-1 for similarities. Only human G-CSF[14] showed significant similarity to BSF-2. Amino acids 28–91 of BSF-2 and amino acids 20–85 of G-CSF match at 17 residues (25.7% of all residues) (Fig. 3*c*). Note that cysteine residues at positions 44, 50, 73 and 83 of BSF-2 match those at positions 39, 45, 67 and 77 of G-CSF exactly. This suggests that BSF-2 is distantly related to G-CSF. If so, intramolecular disulphide bonds would be important in the structure of these proteins.

In Northern blot analysis (Fig. 4) pBSF2.38 DNA hybridized with a single species of messenger RNA ~1,300 nucleotides long which was extracted from activated lymphocytes and TCL-Na1 cells but not from unstimulated control lymphocytes, showing that, like other T-cell-derived lymphokines, BSF-2 mRNA is induced upon mitogenic stimulation.

We have previously reported[10] that cardiac myxoma cells produce high levels of a molecule with BSF-2 activity which was specifically absorbed by the anti-peptide antibody. It has also been shown that bladder cell carcinoma, T24 cells produce a molecule with BSF-2 activity[15]. To show that these cells produce BSF-2 itself, we performed Northern blot analysis. As shown in Fig. 4 (*d*, *e*), BSF-2 mRNA was transcribed in T24 cells and myxoma cells at a much higher level than activated lymphocytes, although the mRNA in myxoma cells was larger than that of lymphocytes. These data indicate that under certain conditions, the BSF-2 gene can be expressed in non-lymphoid tissues. Because patients with cardiac myxoma frequently show several kinds of autoimmune disease-like symptoms[16], it was suggested[10] that deregulated production of BSF-2 might be involved in autoimmune pathogenesis. The fact that BSF-2 gene expression in myxoma cells is high further strengthens this argument.

This work was supported in part by grants from the Ministry of Education, Science and Culture. We thank Dr Edward Barsumian for critical review of the manuscript and Ms K. Kubota, J. Mori and M. Kawata for their secretarial assistance.

Received 4 August; accepted 26 September 1986.

1. Dutton, R. W. *et al. Prog. Immun.* **1**, 355–368 (1971).
2. Schimpl, A. & Wecker, E. *Nature* **237**, 15–17 (1972).
3. Howard, M. & Paul, W. E. *A. Rev. Immun.* **1**, 307–333 (1983).
4. Kishimoto, T. *A. Rev. Immun.* **3**, 135–159 (1985).
5. Okada, M. *et al. J. exp. Med.* **157**, 583–590 (1983).
6. Butler, J. L., Falkoff, R. J. M. & Fauci, A. S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2475–2478 (1984).
7. Hirano, T., Teranishi, T., Lin, B. H. & Onoue, K. *J. Immun.* **133**, 798–802 (1984).
8. Kikutani, H. *et al. J. Immun.* **134**, 990–995 (1985).
9. Hirano, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5490–5494 (1985).
10. Hirano, T. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
11. Shimizu, K. *et al. J. Immun.* **133**, 1728–1733 (1985).
12. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
13. Maxam, A. W. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
14. Nagata, S. *et al. Nature* **319**, 415–418 (1986).
15. Rawle, F. C. *et al. Eur. J. Immun.* **16**, 1017–1019 (1986).
16. Sutton, M. G. S. J., Mercier, L., Giuliani, E. R. & Lie, J. T. *Mayo clin. Proc.* **55**, 371–376 (1980).
17. Hewick, R. M., Hunkapillar, M. W., Hood, L. E. & Dreyer, W. J. *J. biol. Chem.* **256**, 7990–7997 (1981).
18. Tsunasawa, S., Kondo, J. & Sakiyama, F. *J. Biochem., Tokyo* **97**, 701–704 (1985).
19. Taniguchi, T. *et al. Nature* **302**, 305–310 (1983).
20. Mizuuchi, K., Mizuuchi, M. & Gellert, M. *J. molec. Biol.* **156**, 229–243 (1982).
21. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263–269 (1983).
22. Huynh, T. V., Young, R. A. & Davis, R. W. *DNA Cloning* Vol. 1, 49–78 (IRL press, Oxford, 1985).
23. Wood, W. I., Gitshier, J., Lasky, L. A. & Lawn, R. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1585–1588 (1985).
24. Gluzman, Y. *Cell* **23**, 175–182 (1981).
25. Wilbur, W. J. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 726–730 (1983).
26. Lehrach, H., Diamond, D., Wozey, J. M. & Boedtker, H. *Biochemistry* **16**, 4743–4751 (1977).

Intracellularly injected tetanus toxin inhibits exocytosis in bovine adrenal chromaffin cells

Reinhold Penner*, Erwin Neher† & Florian Dreyer*

* Rudolf-Buchheim-Institut für Pharmakologie, Frankfurter Str 107, D-6300 Giessen, FRG
† Max-Planck-Institut für Biophysikalische Chemie, Am Fassberg, D-3400 Göttingen, FRG

The clostridial neurotoxins tetanus and botulinum toxin type A are known to block transmitter release from nerve terminals[1–3], probably by interfering with some essential process controlling exocytosis[3,4] after the entry of Ca^{2+} ions. Although exocytosis occurs in many secretory cells, these toxins show a high specificity for neurones and the secretory response of cultured bovine adrenal medullary cells is not inhibited by exposure to medium containing tetanus or botulinum toxin type A (although it is by botulinum toxin type D)[4]. Here we report that when tetanus toxin and botulinum neurotoxin type A are injected intracellularly into chromaffin cells they strongly inhibit secretion, as revealed by the measurement of cell capacitance[5]. These results indicate that these toxins are normally ineffective in chromaffin cells because they are not bound and internalized, so do not reach their site of action. Furthermore, we have localized the secretion-blocking effects of the toxin to a fragment comprising the light chain covalently linked to part of the heavy chain, suggesting that this part of the molecule contains the active site.

We examined cultured bovine adrenal chromaffin cells, isolated as described[6], using the whole-cell configuration of the patch-clamp technique[7]. This allows one to monitor exocytosis using membrane capacitance as a measure of the cell surface area. Capacitance and surface area change considerably as chromaffin granules fuse with the membrane or endocytotic retrieval of the membrane occurs[5]. Secretion was stimulated by dialysing the cells through a patch-pipette with a solution buffered to a free Ca^{2+} concentration of 1 µM. The total membrane capacitance of control cells (typically 4–8 pF) started to increase immediately upon establishment of the whole-cell configuration, eventually leading to a 2–3-fold capacitance increase within 15–25 min (Fig. 1*a*).

When cells were preloaded with either tetanus toxin or botulinum neurotoxin type A the secretory response to the same challenge (dialysis of cells with Ca^{2+}-containing solution) was effectively removed. In fact, cells pretreated with tetanus toxin consistently showed a slight decrease in total membrane capacitance (Fig. 1*c*), suggesting that in these cells there was more endocytotic than exocytotic activity. This 'reversal effect' was observed only once in cells pretreated with botulinum neurotoxin type A ($n = 10$) which may reflect genuine differences in the ability of the two toxins to inhibit secretion or, as has been suggested[8,9], different sites of action. The typical response to increased intracellular Ca^{2+} levels in cells preloaded with botulinum neurotoxin type A was a slight initial increase in cell

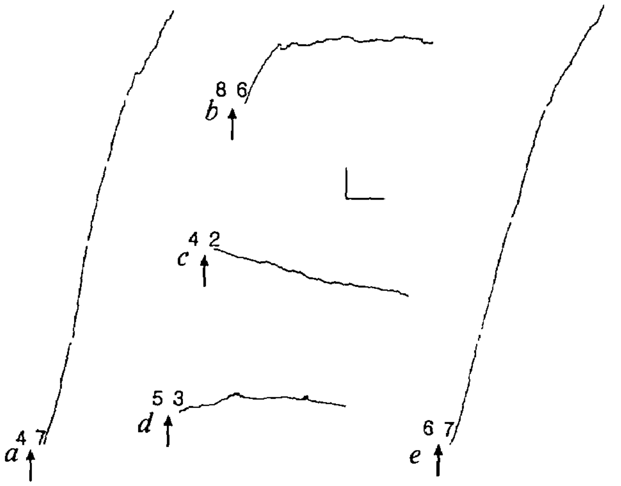


Fig 1 Effects of clostridial toxins on the secretory responses (measured as capacitance changes) of bovine chromaffin cells stimulated by elevation of the intracellular Ca^{2+} concentration *a*, Typical time course of capacitance changes recorded from a stimulated control cell *b*, Typical inhibitory effect of botulinum A toxin on the response of a stimulated cell *c*, Inhibition of exocytosis by tetanus toxin *d*, Inhibition of exocytosis in cells preloaded with fragment B at 25 μg ml^{-1} in the pipette *e*, Capacitance increase recorded from a cell preloaded with fragment C at 25 μg ml^{-1} in the pipette Capacitance measurements were performed according to the method described by Neher and Marty[5] Briefly, adrenal medullary chromaffin cells isolated as described[6] and kept in short-term culture (1–5 days) were voltage clamped at −70 mV and subjected to a sinusoidal voltage command of 18 mV r m s at 795 Hz The resulting current was fed into a lock-in amplifier which delivered the capacitive component of the signal when tuned to the desired phase angle with respect to the reference signal For the large capacitance changes of the control and fragment C traces, these had to be reconstructed from several segments (interruptions) since the lock-in measurement requires recompensation after large deflections[5] In all traces establishment of the whole-cell recording configuration is indicated by arrows Numbers given at the start of each trace correspond to initial cell capacitance values (in pF) as determined by the dial setting of the EPC-7 amplifier (List Electronic, Darmstadt) Average capacitance values were 5 9±0 3 pF (mean±s e m, $n=25$), series resistance values, as read from the EPC-7 settings, were 3 9±0 5 MΩ *a*, To stimulate secretion a pipette filling solution of the following composition was used (mM) K glutamate 135, NaCl 20, $MgCl_2$ 4, ATP 0 5, EGTA 10, $CaCl_2$ 9 (free Ca^{2+} 1 μM), HEPES 10 (*p*H 7 2) The bath solution contained (mM) NaCl 140, KCl 2 8, $MgCl_2$ 2, $CaCl_2$ 2, HEPES 10 (*p*H 7 2) The cell underwent the pretreatments described in *b* except that toxin was omitted from the pipette when 'loading' the cell *b*, The cell was loaded with the toxin in the whole-cell configuration for 3 min and incubated for 1 h at 37 °C Botulinum toxin ($LD_{50}=8$ ng kg^{-1}, mouse s c) was used at a concentration of 50 μg ml^{-1} in a pipette filling solution containing (mM) K glutamate 135, NaCl 20, $MgCl_2$ 4, EGTA 1, HEPES 10 (*p*H 7 2) Both the procedure of loading and the actual test for the secretory response were carried out at 30 °C *c*, Cells were pretreated, incubated and challenged as in *b* except when loading the cells 50 μg ml^{-1} tetanus toxin ($LD_{50}=3$ ng kg^{-1}, mouse s c) was present in the electrode filling solution Vertical scale bar, 0 2 pF, horizontal scale bar, 1 min

capacitance, followed by a phase in which no major alteration was seen (Fig 1*b*) Cells preloaded using pipettes containing no toxin or denatured (boiled) toxins showed a normal secretory response The rates of release resulting from an elevation of the intracellular Ca^{2+} of control and toxin-treated cells are shown in Fig 2

Both secretion and tetanus toxin effects are temperature dependent Lowering the temperature slows, and can even prevent, the inhibitory effects of tetanus toxin at nerve terminals[10] We therefore carried out experiments at different temperatures Whereas secretion at 30 °C proceeded at an average rate of 19 7% per min (see Fig 2), at 20 °C it was reduced to 7 5 ± 1 4% per min (mean ± s e m, $n=6$) At 20 °C the release rate of cells pretreated with tetanus toxin (5 7 ± 1 5% per min, mean ± s e m, $n=11$) was not significantly different from that of control cells This finding clearly indicates that the temperature-dependent step in the action of tetanus toxin is subsequent to binding and internalization Combined with the data obtained using boiled toxins, these results strongly support the conclusion that the inhibitory actions of tetanus and botulinum A toxin on the release mechanism are highly specific and not simply due to the intrusion of a protein of high relative molecular mass (M_r) into the cell

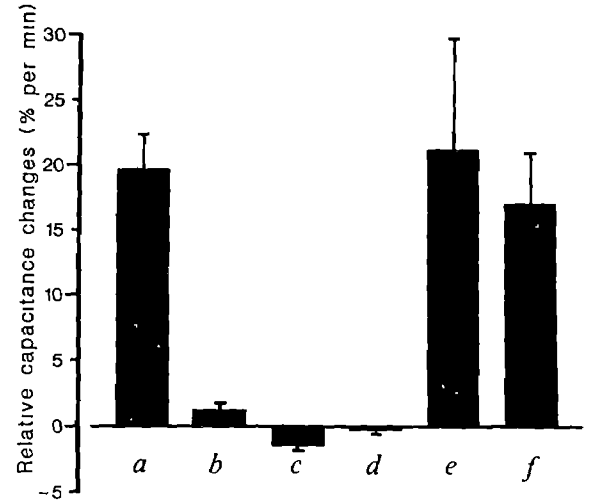


Fig 2 Effects of clostridial toxins on secretion from bovine adrenal medullary cells *a*, Control (mean ± s e m of 14 cells), *b*, botulinum toxin A (10 cells), *c*, tetanus toxin (5 cells), *d*, fragment B (7 cells), *e*, fragment C (4 cells), *f*, denatured tetanus toxin (3 cells) The measure of release chosen was the relative capacitance increase (% of the initial value) per unit time (min) Care was taken to assure relatively uniform cell sizes so that differences in initial capacitance values for each set of experiments were small To determine the release rates capacitance values were measured 2–4 min after transition from cell-attached to whole-cell configuration which usually corresponded to the time in which the capacitance increase was linear (see Fig 1*a* and *e*)

We used the same system to test for intracellular effects of fragments B and C of tetanus toxin Fragment C is a part of the toxin's heavy chain of M_r 50,000 which is thought to mediate binding of the toxin to cell membranes[11,12] Fragment B consists of the light chain covalently linked to the remaining part of the heavy chain (M_r 100,000) These fragments are obtained by enzymatic treatment with papain, and are non-toxic at doses of 10 mg kg^{-1} when injected into mice (the LD_{50} for tetanus toxin is 3 ng kg^{-1}, mouse subcutaneous (s c)[13] Fragment C did not inhibit the responses of chromaffin cells to increased intracellular Ca^{2+} levels (Figs 1*e* and 2), but fragment B effectively suppressed the secretory response (see Figs 1*d* and 2), suggesting that this fragment bears the active site for blocking exocytosis We have thus shown that a fragment of tetanus toxin that is otherwise non-toxic is able to block secretion

It is not known whether fragmentation of the toxin into B and C occurs fortuitiously or whether it is integral to the inhibitory actions of tetanus toxin It appears that some fragmentation occurs *in vivo*, but most of the toxin extracted from the spinal cord (the destination of toxin after uptake into motor nerve terminals and retrograde transport) is indistinguishable from native toxin[14] This suggests that the proposed uptake mechanism of clostridial neurotoxins[2] (drawn by analogy with diptheria toxin and Semliki Forest virus) by receptor-mediated endocytosis followed by expulsion of the toxin's light chain through a 'channel' formed by the toxin's heavy chain into the cytoplasm may not apply From our method of toxin administration it is clear that, even if this described uptake mechanism is used, any specific structural changes occurring during uptake are not absolutely required for the actions of the toxins If changes that are required for activity occur, our results suggest that they are more likely to happen within the cytoplasm

These results support the hypothesis that tetanus and botulinum A toxin act beyond the rise of intracellular Ca^{2+}

concentration, interfering with an as yet unknown step late in the chain of events leading to exocytosis. Our data show that this crucial step is not restricted to the release process from nerve terminals, but may be quite generally associated with vesicular release. Clostridial neurotoxins may serve as valuable tools in studying late steps in exocytosis. The whole-cell recording technique combined with capacitance measurement is useful because it allows the toxins to reach their site of action directly in cells which lack the prerequisite membrane binding and internalization features and allows tests of fragments that are not normally internalized. Experiments using smaller fragments may allow the characterization of a 'minimal toxin' small enough to be introduced into high-voltage[15] or drug-permeabilized[16,17] cells.

We thank Mr U. Weller and Professor E. Habermann for supplying purified fragments B and C. Tetanus toxin was from the Behringwerke (Marburg, FRG), botulinum toxin type A was a gift from Dr Schantz (Food Research Institute, Madison). We also thank Dr S. DeRiemer for valuable comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 47).

Received 7 July; accepted 1 September 1986.

1. Mellanby, J. *J. Neurosci.* **11**, 29–34 (1984).
2. Simpson, L. L. *A. Rev. Pharmac. Tox.* **26**, 427–453 (1986).
3. Habermann, E. & Dreyer, F. *Curr. Top. Microbiol. Immun.* **129**, 93–179 (1986).
4. Knight, D. E., Tonge, D. A. & Baker, P. F. *Nature* **317**, 719–721 (1985).
5. Neher, E. & Marty, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6712–6716 (1982).
6. Fenwick, E. M., Fajdiga, P. B., Howe, N. B. S. & Livett, B. G. *J. Cell Biol.* **76**, 12–30 (1978).
7. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
8. Dreyer, F. & Schmitt, A. *Neurosci. Lett.* **26**, 307–311 (1981).
9. Dreyer, F. *et al. J. Physiol. (Paris)* **79**, 252–258 (1984).
10. Schmitt, A., Dreyer, F. & John, Chr. *Naunyn Schmiedebergs Archs Pharmak.* **317**, 326–330 (1981).
11. Morris, N. P. *et al. J. biol. Chem.* **255**, 6071–6076 (1980).
12. Goldberg, R. L., Costa, T., Habig, W. H., Kohn, L. D. & Hardegree, M. C. *Molec. Pharmac.* **20**, 565–570 (1981).
13. Weller, U., Taylor, C. F. & Habermann, E. *Toxicon* (in the press).
14. Habermann, E., Wellhöner, H. H. & Raker, K. O. *Naunyn Schmiedebergs Archs Pharmak.* **299**, 187–196 (1977).
15. Baker, P. F. & Knight, D. E. *Nature* **276**, 620–622 (1978).
16. Dunn, L. A. & Holz, R. W. *J. biol. Chem.* **258**, 4989–4993 (1983).
17. Wilson, S. P. & Kirshner, N. *J. biol. Chem.* **258**, 4994–5000 (1983).

The essential light chains constitute part of the active site of smooth muscle myosin

Yoh Okamoto*, Takamitsu Sekine*, Jean Grammer† & Ralph G. Yount†

* Department of Biochemistry, Juntendo University School of Medicine, Tokyo, Japan
† Biochemistry/Biophysics Program, Institute of Biological Chemistry and Department of Chemistry, Washington State University, Pullman, Washington 99164-4630, USA

Myosin, a major contractile protein, characteristically possesses a long coiled-coil α-helical tail and two heads. Each head contains both an actin binding site and an ATPase site and is formed from the NH_2-terminal half of one of the two heavy chains (relative molecular mass, M_r, 200,000) and a pair of light chains; the so-called regulatory and essential light chains of approximately M_r 20,000 each. Recently we have identified[1] Trp 130 of the myosin heavy chain from rabbit skeletal muscle as an active-site amino-acid residue after labelling with a new photoaffinity analogue of ADP, N-(4-azido-2-nitrophenyl)-2-aminoethyl diphosphate (NANDP)[2]. Nonspecific labelling was eliminated by first trapping NANDP at the active site with thiol crosslinking agents[3]. Exclusive labelling of the heavy chains with no labelling of the light chains agreed with previous findings[4,5] that the heavy chains alone contain the actin-activated Mg-ATPase activity of rabbit skeletal myosin. Here we report similar photolabelling experiments with smooth muscle myosin (chicken gizzard) in which ^{3}H-NANDP is trapped at the active site with vanadate[6] and which show that both the heavy chains and the essential light chains are labelled. The results indicate that both chains contribute to the ATP binding site and represent the first direct evidence for participation of the essential light chains in the active site of any type of myosin.

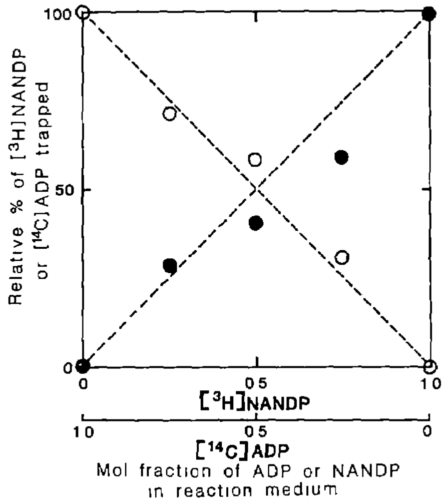


Fig. 1 The relative loss of Ca^{2+} and K^+-EDTA ATPase activities of gizzard myosin as a function of vanadate-trapped [^{3}H]NANDP. Gizzard myosin (4 mg ml^{-1}) in 0.5 M KCl, 2.0 mM $MgCl_2$, 20 mM Tris-Cl, pH 8.0, was incubated with 0.7 mM vanadate and varying concentrations of [^{3}H]NANDP[1,2] for 2 h at 0 °C. Aliquots (0.9 ml) were passed over Sephadex G-25 (PD-10) columns to remove vanadate and untrapped [^{3}H]NANDP. The peak containing myosin-[^{3}H]NANDP-vanadate was assayed for protein (biuret method) and for Ca^{2+} (○) and K^+-EDTA (●) ATPase activities[18]. The [^{3}H]NANDP content was determined by scintillation counting. Actin-free gizzard myosin was prepared from fresh chicken gizzards essentially by the procedure of Ebashi[30] and was in the dephosphorylated form. In addition, all experiments were performed at high ionic strength where all the myosin molecules are in the 6S unfolded state[16]. Vanadate solutions were prepared and analysed according to Goodno[31].

It is of primary importance in photoaffinity labelling studies to eliminate nonspecific labelling. The discoveries that nucleotides and nucleotide analogues could be trapped in the active site of myosin either by crosslinking two reactive thiols, SH_1 and SH_2 (refs 3, 7) or by forming a stable nucleotide complex with vanadate[6] have given highly effective ways to achieve specific photolabelling. Both approaches lead to an ~10^4 decrease in the off-rate of bound nucleotide and give complexes with lifetimes of several days at 0 °C. It is likely that both types of trapping yield quite similar active-site structures since we have found that binding of NANDP can be stabilized at the active site of cardiac myosin by either method (Y.O. and R.G.Y., manuscript in preparation). Photolysis in each case resulted in 50–60% covalent photolabelling of a heavy-chain peptide which was almost identical in sequence to that obtained from rabbit skeletal myosin[1]. Similar studies with other ATP photoaffinity analogues all indicate that heavy chains of skeletal myosin and not light chains are modified[7,8]. Earlier studies by Szilagyi *et al*[9] with an analogue containing a nitroarylazide group on the ribose ring of ATP also photolabelled predominantly the heavy chains although a small amount of labelling on the light chains could not be ruled out. These studies did not, however, utilize either of the two trapping methods mentioned above.

It was of interest to extend our studies to smooth muscle myosin because these molecules, unlike myosins from skeletal and cardiac muscles, exhibit myosin-linked regulation of contraction[10,11]. In addition, gizzard myosin is known to undergo

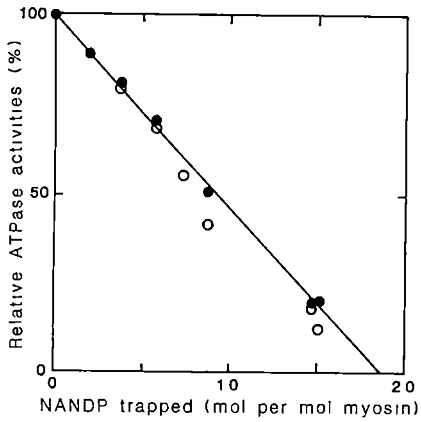


Fig 2 Competitive trapping of [^{14}C]ADP (O) or [^{3}H]NANDP (●) on gizzard myosin by use of the vanadate ion Myosin (4 mg ml^{-1}) in 0 5 M KCl, 2 0 mM MgCl$_2$, 20 mM Tris-Cl, *p*H 8 0, was incubated with 1 mM vanadate and varying ratios of [^{14}C]ADP (Amersham) to [^{3}H]NANDP (total concentration of ADP plus NANDP=0 1 mM) for 2 h at 0 °C The myosin-[^{14}C]ADP (or [^{3}H]NANDP)-vanadate complexes were purified and analysed as described in Fig 1 The maximal amount of trapping under these conditions was 1 34 mol [^{3}H]NANDP per mol myosin and 1 42 mol [^{14}C]ADP per mol myosin

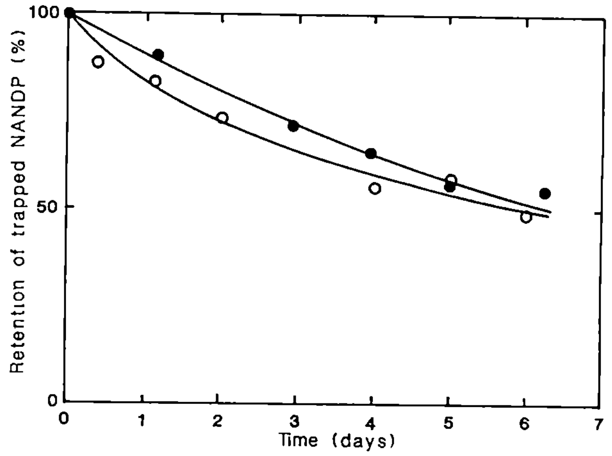


Fig 3 Relative stability of [^{3}H]NANDP-vanadate complexes of gizzard (O) and skeletal (●) myosins The vanadate-[^{3}H]NANDP complexes of gizzard and skeletal myosin were prepared and purified as described in Fig 1 and stored at 0 °C At appropriate times, 0 1-ml aliquots were passed over PD-10 columns (Pharmacia) and analysed for radioactivity and protein concentration[1,18]

large changes in conformation as a function of phosphorylation of the regulatory light chain (LC$_{20}$) (refs 12–14), the binding of Mg-ATP[15], or changes in ionic strength[16,17] In particular, it seemed possible that specific photoprobes could reveal changes in the composition of the active site as a function of these induced conformational changes Moreover, earlier studies using modified gizzard myosin[18] had indicated a close connection between ATP cleavage and changes in a fluorescently-labelled essential light chain, suggesting that essential light chains (LC$_{17}$) were closely linked to the active site We trapped [^{3}H]NANDP on gizzard myosin with a vanadate ion in a manner similar to that described by Goodno for ADP on skeletal myosin[6] In this work, as shown in Fig 1, vanadate stabilized the binding of [^{3}H]NANDP to the active site of gizzard myosin as judged by the parallel loss of both the Ca^{2+}- and K^{+}-EDTA ATPase activities and by the nearly stoichiometric trapping of [^{3}H]NANDP to each head

Competitive vanadate trapping experiments were done with [^{14}C]ADP and [^{3}H]NANDP to verify further that NANDP was binding at the active site Figure 2 shows that as the mol fraction of [^{14}C]ADP was increased relative to the mol fraction of [^{3}H]NANDP, more [^{14}C]ADP was trapped but that the total amount of diphosphate compounds trapped remained constant These results indicate that [^{3}H]NANDP and [^{14}C]ADP bind with comparable affinities to a single type of site and that they compete effectively with each other for stable binding to gizzard myosin

The stability of vanadate trapped [^{3}H]NANDP on gizzard myosin was of considerable interest since specific photoincorporation of NANDP into skeletal myosin has been shown to require previous trapping of some type[1,2] As Fig 3 illustrates, both the gizzard myosin and the skeletal myosin-[^{3}H]NANDP-vanadate complexes show essentially identical long-term stabilities In both cases the half-life at 0 °C was greater than 5 days indicating that leakage of [^{3}H]NANDP from the active site would be negligible during the short time required for purification and photolysis of the trapped complex It should be noted that care was taken to remove traces of actin from gizzard myosin and to keep the temperature near 0 °C Both factors, F-actin and elevated temperatures, are known to destabilize nucleotides trapped at the active site of skeletal myosin[7,19]

The gizzard myosin-[^{3}H]NANDP-vanadate complex purified by gel filtration was photolysed as described in Fig 4 Approximately 0 20 mol per mol active site or 30% of the trapped [^{3}H]NANDP (0 68 mol per mol active site) was found to be attached covalently after 40 min of photolysis It was found to be important in the photolabelling to use irradiation above 400 nm as ultraviolet light caused a photochemical reaction of the trapped vanadate with myosin which in turn allowed [^{3}H]NANDP (or in other experiments [^{3}H]ADP) to leak out of the active site (J G , R G Y and Y O , unpublished results) In control experiments using the conditions in Fig 4, less than 2% of either [^{3}H]NANDP or [^{3}H]ADP trapped with vanadate on gizzard myosin was lost in 40 min of irradiation Complete resolution of the heavy chains from the two light chains, LC$_{20}$ (regulatory) and LC$_{17}$ (essential), was accomplished using high-speed gel filtration on TSK-SW columns in the presence of SDS[20] Surprisingly, as Fig 4 illustrates, some 20% of the photoincorporated [^{3}H]NANDP was associated with LC$_{17}$, with the remainder of the labelling occurring with the heavy chains The LC$_{20}$ chains were not labelled These results suggest that parts of both LC$_{17}$ and the heavy chain combine to form the purine-binding site for ATP Such a finding is reminiscent of the structure of the immunoglobulin G molecule, in which parts of both the heavy chains and the light chains function to form the antigen binding site[21] For IgG, as appears to be the case here, affinity labels gave the first indication that both chains were necessary to form the complete binding site[22,23]

The validity of the interpretation of our results is supported further by the remarkable similarity between NANTP and ATP in their ability to support tension development and relaxation of chemically skinned muscle fibres from all three major types of muscle (R Bridenbaugh, K Nakamaye, G Kerrick and R Y, in preparation) In addition, support of effective tension development in smooth muscle by NANTP required previous thiophosphorylation[24] of LC$_{20}$ because NANTP was not an effective substrate for myosin light-chain kinase These results are clear evidence that NANTP and, by analogy, NANDP, binds to the active site of gizzard myosin in a manner similar if not identical to ATP

The marked difference in the labelling pattern observed for smooth muscle myosin compared with that seen with striated muscle myosins is surprising It is most likely that this difference reflects variations in the heavy-chain sequences because the sequences of the essential light chains from chicken gizzard, heart and skeletal muscle are all highly homologous[25] Currently, little is known about the sequence of smooth muscle myosin

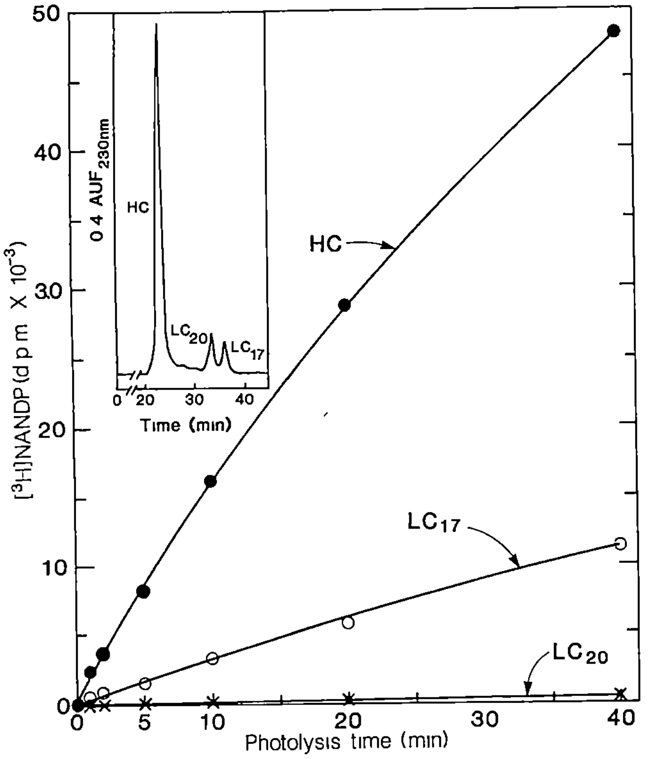


Fig 4 The subunit location of [^{3}H]NANDP photoincorporated into gizzard myosin The gizzard myosin-[^{3}H]NANDP-vanadate complex was prepared essentially as described in Fig 1 except trapping was done overnight using 1 mM Na vanadate, 25 μM [^{3}H]NANDP and 3 7 mg ml^{-1} gizzard myosin The Sephadex G-25 purified complex (10 ml) was irradiated in a 6 cm i d Petri dish at 0 °C with a Ushio (100 W) high-pressure mercury lamp The Petri dish was encased in aluminium foil and the top covered with a glass filter (L-42, Toshiba) to remove irradiation below 400 nm At appropriate irradiation times, samples were taken from the Petri dish, cold trichloroacetic acid added to 5% and the protein precipitate collected by centrifugation The precipitate was dissolved in Tris buffer and SDS (20% solution) added to 0 2% The solution was heated at 100 °C for 3 min, classified by centrifugation and 100 μl of the supernatant passed over serially linked TSK G4000SW and TSK G3000SW columns (7 5 mm × 60 cm each, Toyo Soda) using Beckman-Altex pumps at 100 p s i, flow rate = 1 ml per min All three myosin polypeptides were well resolved (see inset) with the peak of the heavy chain (HC) eluting at 23 min, LC_{20} at 33 min and LC_{17} at 36 min The counts were normalized at each time point to 1 4 mg of myosin (~6 0 nmol of active sites) The distribution of radioactivity was 78±4% into HC and 17±2% into LC_{17}

heavy chains It may be only in the case of myosin-linked regulated systems[26–29] that the essential light chains function to provide part of the ATP binding site Experiments to test this suggestion and to test the possible effect of phosphorylation of LC_{20} and the folding of the heavy chain on the subunit labelling pattern of gizzard myosin are currently under way

This research was supported in part by a research grant from the Ministry of Education, Science and Culture of Japan to Y O and T S and by NIH grant AM-05195 and by the Muscular Dystrophy Association to R G Y

Received 11 June, accepted 5 August 1986

1 Okamoto, Y & Yount, R. G *Proc natn Acad Sci USA* **82**, 1575-1579 (1985)
2 Nakamaye, K, Wells, J A, Bridenbaugh, R, Okamoto, Y & Yount, R G *Biochemistry* **24**, 5226-5235 (1985)
3 Wells, J A & Yount, R. G *Proc natn Acad Sci USA* **76**, 4966-4970 (1979)
4 Wagner, P D & Giniger, E *Nature* **292**, 560-562 (1981)
5 Sivaramakrishnan, M & Burke, M *J biol Chem* **257**, 1102-1105 (1982)
6 Goodno, C *Proc natn Acad Sci USA* **76**, 2620-2624 (1979)
7 Wells, J A & Yount, R G *Meth Enzym* **85**, 95-116 (1982)
8 Mahmood, R & Yount, R. G *J biol Chem* **259**, 12956-12959 (1984)
9 Szilagyi, K, Balint, M, Sreter, F A & Gergely, J *Biochem biophys Res Commun* **87**, 936-945 (1979)
10 Adelstein, R S & Eisenberg, E *A Rev Biochem* **49**, 921-956 (1980)
11 Marston, S *Prog Biophys molec Biol* **41**, 1-41 (1982)
12 Onishi, H & Wakabayashi, T *J Biochem (Tokyo)* **92**, 871-879 (1982)
13 Trybus, K M, Huiatt, T W & Lowey, S *Proc natn Acad Sci USA* **79**, 6151-6155 (1982)
14 Craig, R., Smith, R & Kendrick Jones, J *Nature* **302**, 436-439 (1983)
15 Suzuki, H, Onishi, H, Takahashi, K. & Watanabe, S *J Biochem (Tokyo)* **84**, 1529-1542 (1978)
16 Ikebe, M, Hinkins, S & Hartshorne, D J *Biochemistry* **22**, 4580-4587 (1983)
17 Trybus, K M & Lowey, S *J biol Chem* **259**, 8564-8571 (1984)
18 Okamoto, Y & Sekine, T *J Biochem (Tokyo)* **87**, 167-178 (1980)
19 Goodno, C C & Taylor, E W *Proc natn Acad Sci USA* **79**, 21-25 (1982)
20 Kato, Y, Komiya, K, Sasaki, H & Hashimoto, T *J Chromatogr* **193**, 29-36 (1980)
21 Poljak, R J *Nature* **256**, 373-376 (1975)
22 Singer, S J & Doolittle, R F *Science* **153**, 13-25 (1968)
23 Good, A H, Ovary, E & Singer, S J *Biochemistry* **7**, 1304-1310 (1968)
24 Cassidy, P, Hoar, P E & Kerrick, W G L *J biol Chem* **254**, 11148-11153 (1979)
25 Matsuda, G *Adv Biophys* **16**, 185-218 (1983)
26 Adelstein, R S & Conti, M A *Nature* **256**, 597-598 (1975)
27 Scordilis, S P & Adelstein, R S *Nature* **268**, 558-560 (1977)
28 Trotter, J A & Adelstein, R S *J biol Chem* **254**, 8781-8785 (1979)
29 Szent-Gyorgyi, A G, Szentkiralyi, E. M & Kendrick-Jones, J *J molec Biol* **74**, 179-203 (1973)
30 Ebashi, S *J Biochem* **79**, 229-231 (1976)
31 Goodno, C *Meth Enzym* **85**, 116-123 (1982)

Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases

Peter K. Hwang & Robert J. Fletterick

Department of Biochemistry and Biophysics,
University of California, San Francisco, California 94143, USA

The activity of many proteins in eukaryotic cells is regulated by reversible covalent phosphorylation[1]. This regulatory modification is often linked to other allosteric controls within the same protein[2], and such overlapping regulatory mechanisms are best characterized for glycogen phosphorylase (EC 2.4.1.1). Phosphorylases from different organisms or cell types exhibit markedly contrasting regulatory features[3]; this makes the enzyme attractive for studying the evolution of interacting molecular regulatory mechanisms[4,5]. Extensive biochemical and crystallographic studies of rabbit muscle phosphorylase have led to a characterization of five regulatory regions (phosphorylation, glycogen storage, AMP, glucose and purine sites)[6–8]. Here we report the complete primary structure of the yeast *Saccharomyces cerevisiae* glycogen phosphorylase, deduced from the sequence of the cloned gene. Regions that are highly conserved between muscle and yeast enzymes include the active site, the glycogen storage site and possibly the glucose and purine inhibition sites. Partial conservation of the residues involved in AMP-binding suggests a binding site for the yeast enzyme inhibitor, glucose 6-phosphate[9,10]. Other parts of the AMP site and the intersubunit contacts involved in AMP allostery are disrupted in the yeast enzyme by extreme sequence divergence. The poor alignment of amino termini and lack of homology at phosphorylation sites indicate that regulation by reversible phosphorylation evolved independently in yeast and vertebrate phosphorylases.

The yeast phosphorylase gene was isolated from a collection of yeast genomic DNA fragments cloned into plasmid vectors (see Fig 1 legend) An 890-residue amino-acid sequence was deduced from the nucleotide sequence of a 2 7-kilobase (kb) uninterrupted open reading frame We have compared the yeast amino-acid sequence to that of rabbit muscle phosphorylase (Fig 2) The best alignment requires the introduction of 14 insertions or deletions in one sequence relative to the other (Fig 2 legend) Although the alignment of N termini is somewhat arbitrary, as the two sequences differ considerably, the overall similarity between the yeast and muscle protein sequences is quite high The yeast enzyme exhibits 49% amino-acid identity

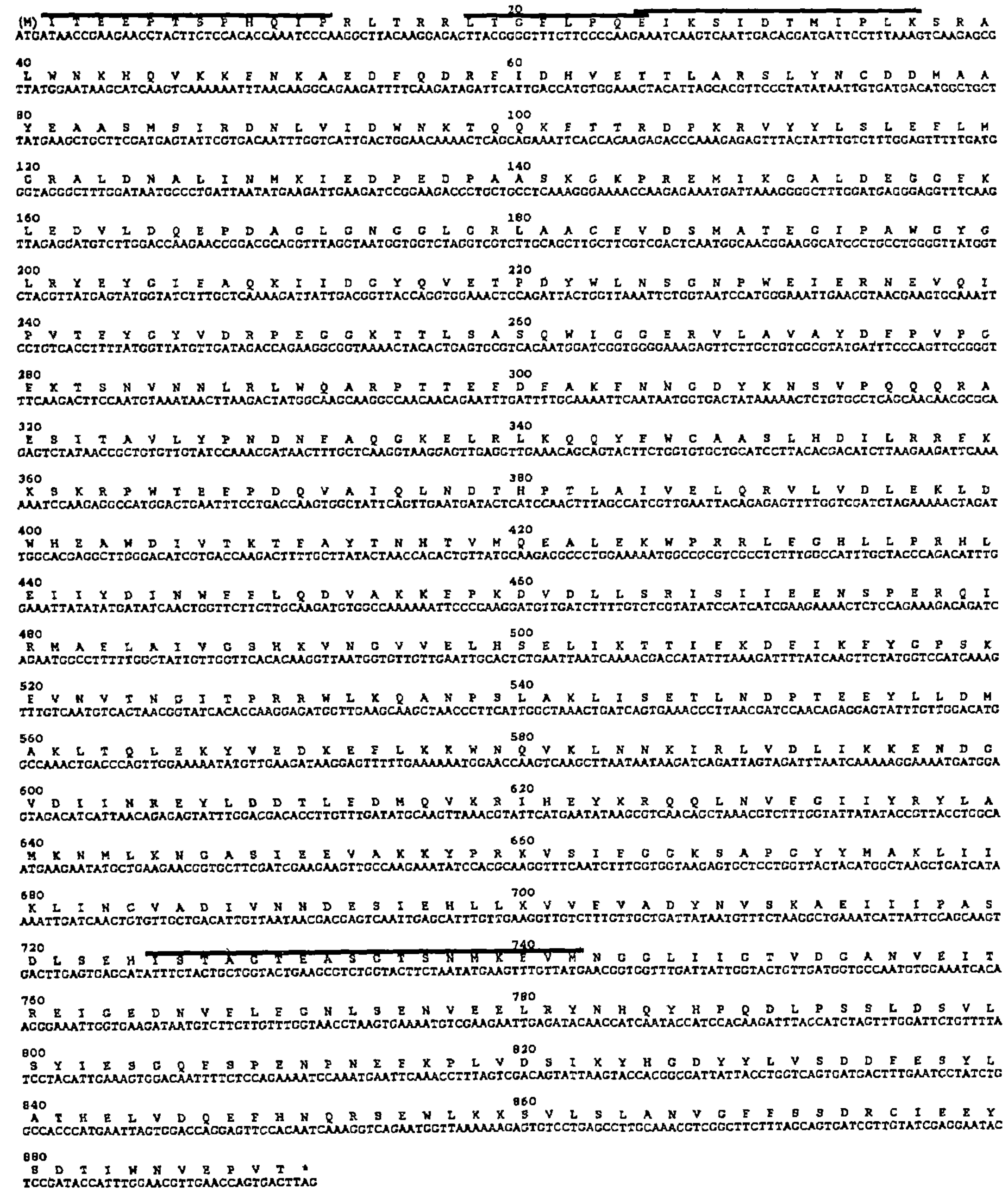


Fig. 1 Nucleotide sequence and derived amino-acid sequence of the glycogen phosphorylase gene in *S. cerevisiae*. The base composition of the DNA sequence is 59.8% A T. The deduced amino-acid composition matches (within 5% for all residues) the previously determined amino-acid analysis[9]. The calculated relative molecular mass of the deduced protein sequence, 102,400 (102,700 if the pyridoxal phosphate is included), agrees with experimentally determined values[9,19]. The predicted C-terminus residue, threonine, also agrees with the result obtained for carboxypeptidase-treated enzyme[19]. Bars, amino-acid sequences previously determined by peptide sequencing: (1) residues 1-12 (E. Schiltz, personal communication), (2) residues 18-25 (ref. 18), (3) residues 25-36 (ref. 19), (4) residues 725-742 (ref. 18).

Methods. The amino-acid sequence of a pyridoxal phosphate-bound peptide of yeast phosphorylase[18] was used to design a mixed oligonucleotide probe (a 32-fold degenerate 18-mer encoding the peptide sequence N-M-K-F-V-M). The probe was radiolabelled by kinase reaction with [γ-^{32}P]ATP and used to screen several yeast genomic DNA libraries contained in plasmid vectors (M. Rose, P. Novick and J. Thomas, personal communication, ref. 30). Hybridizations were carried out in 2×SSC, 0.1% SDS, at 30 °C for 1 h, with $\sim 10^6$ c.p.m. ml^{-1} of probe; washes for 10-15 min were carried out under similar conditions. Pools of plasmids that hybridized to the probe were used to transform *Escherichia coli*. Clones containing putative phosphorylase sequences were subjected to restriction analysis. Fragments containing the oligonucleotide sequence were identified by Southern blot hybridization[31] and sequenced by the M13/dideoxy method[32]. Approximately 92% of the nucleotide sequence was determined in both directions or by multiple sequencing.

to the rabbit muscle enzyme over alignable sequences (403 identical residues at 820 positions). Including conservative replacements at 19% of the aligned positions, the similarity between the yeast and rabbit muscle enzymes is 68%.

Despite the large number of alignment gaps, in most places yeast phosphorylase probably shares a similar tertiary structure with the muscle enzyme. First, all the gaps map to positions on the surface of the rabbit muscle phosphorylase three-dimensional structure and could easily be accommodated in the same structure. Second, most of the non-conservative amino-acid substitutions between the two sequences involve surface residues of the muscle enzyme.

In the catalytic region, all residues involved in cofactor binding or substrate binding are conserved. At the glycogen storage site, only three of the eight residues involved in substrate binding are different in the yeast enzyme, and two of these changes are conservative (Fig. 2).

Two of the five regulatory sites in rabbit muscle phosphorylase bind inhibitors. The enzyme is synergistically inhibited by glucose and purines[11], both of which bind in the active-site cleft formed between the N- and C-terminal domains[12-14]. Glucose binds directly at the active site loop (to Asn 284), which is conserved in the yeast enzyme (Fig. 3). Asp 283, which forms an important part of the loop interaction with Arg 569, is also conserved. Purines bind in a hydrophobic pocket formed by Phe 285 and Tyr 613; these residues are also conserved in the yeast sequence. Conservation of these glucose-binding and purine-binding residues suggests that glucose and purines are involved in the regulation of yeast phosphorylase although this has not previously been investigated.

```
YEAST            ITEEPTSPH QIPRLTRRLT GFLPQEIKSI DTMIPLKSRA LWNKHQVKKF NKAEDFQDRF
RABBIT          ---------- ---------- ---------S RPLSDQEK K QISVRGLAGV ENVTELKKN
                                                     d  ddd ddpddd           d  d
                                                                   20

YEAST           IDHVETTLAR SLYNCDDMVA YEAASMSIRD NLVIDWNKTQ QKFTTRDPKR VYYLSLEFLM
RABBIT          NR LHF  VK DRNVATPRDY  F LAHTV   H  GR IR    HYYEK     I       Y
                dd ddd     apaa     d         dd   d  ddpa  a  a
                       40                    60                    80

YEAST           GRALDNALIN MKIEDPEDPA ASKGKPREMI KGALDEGGFK LEDVLDQEPD AGLGNGGLGR
RABBIT            T Q TMV  LAL NAC E  TY-------- -----QL LD M ELEEI E
                                       d              dddd                   gg v
                      100                                  120

YEAST           LAACFVDSMA TEGIPAWGYG LRYEYGIFAQ KIIDGYQVET PDYWLNSGNP WEIERNEVQI
RABBIT               L      L LA Y    I   F   N    CG W M E A D  RY      KA P FTL
                                         dd        d    ddd d d   d      dddddd
               140                   160                   180

YEAST           PVTFYGYVDR PEGGKTTLSA SQWIGGERVL AVAYDEPVPG FKTSNVNNLR LWQARPTTEF
RABBIT            H   R EH TSQ AK---- -- VDTQV    MP  T     YRNNV  TM    S KAPND
                                                                    a      d  d d
               200                         220                   240

YEAST           DLNKFNNGDY KNSVAQQQRA ESITAVLYPN DNFAQGKELR LKQQYFWCAA SLHDILRRFK
RABBIT          N KD  V G  IQA LDRNL   N SR         FE         E  VV   T Q  I
                d d      d    d  d             c ggc                         aa
                     260                   280                   300

YEAST           KSK------- -RPWTEFPDQ VAIQLNDTHP TLAIVELQRV LVDLEKLDWH EAWDIVTKTF
RABBIT          S  FGCRDPV RTNFDA   K            S   P  M        R   D K  EVTV  C
                     320                   340                   360

YEAST           AYTNHTVMQE ALEKWPRRLF GHLLPRHLEI IYDINWFFLE DVAKKFPKDV DLLSRISIIE
RABBIT                 LP     R  VH L ET      Q    E  QR  N R  AA  G    R R M LV
                    gg                        s   s  s   s                s  s s
                     380                   400                   420

YEAST           ENSPERQIRM AFLAINGSHK VNGVVELHSE LIKTTIFKDF IKFYGPSKFV NVTNGITPRR
RABBIT           GAVK I-N   H C A   A     ARI    IL K       YELE- HK Q  K
                s                        g                                g v
                      440                   460                    480

YEAST           WLKQANPSLA KLISETLNDP TEEYLLDMAK LTQLEKYVED KEFLKKWNQY KLNNKIRLVD
RABBIT            VLC  G   EI A RIG-- -   IS LDQ  RK LS  D  EA IRDVAKV  QE  LKFAA
                v
                       500                       520                   540

YEAST           LIKKENDGVD IINREYLDDT LFDMQVKRIH EYKRQQLNVF GIIYRYLAMK NMLKNGASIE
RABBIT          YLER YKVHI N------PNS    V            L  CL HV TL NRI  K---------
                                            gg c  gg
                                 560                   580

YEAST           EVARKYPRKV SIFGGKSAPG YYMAKLIIKL INCVADIVNN DESIEHLLKV VFVADYNVSK
RABBIT           PNKFVVPRT VMI   A     H   M      TAIG V  H  PVVGDR R  I LEN R  L
                                      c                                     v v
                     600                   620                   640

YEAST           AEIIIPASDL SEHISTAGTE ASGTSNMKFV MNGGLIIGTV DGANVEITRE IGEDNVFLFG
RABBIT            KV   A     Q            G    M L  A T   M       MAE  A  E F I
                                    g gg gg vv   vv
                     660                   680                   700

YEAST           NLSENVEELR ---YNHQYHP QDLPSSLDSV LSYIESGQFS PENPNEFKPL VDSIKYHGDY
RABBIT          MRV D DR D QRG  A EYY DRI -E RQI IEQLS  F    KQ DL  DI  NMLM- H R
                     720                     740                  760

YEAST           YLVSDDFESY LATHELVDQE FHNQRSEWLK KSVLSLANVG FFSSDRCIEE YSDTIWNVEP
RABBIT          FK FA Y E  VKCQ R SAL YK P -  TR MVIRNI TS  K S   T AQ  ARE  G
                      780                      800                   820

YEAST           VT*
RABBIT          SRQRLPAPD EKIP*
                         840
```

Fig 2 Comparison of the primary structures of yeast and rabbit muscle phosphorylases The rabbit muscle phosphorylase sequence includes the corrections of Nakano *et al*[33] and has been renumbered accordingly The Dayhoff alignment program[34] was used to obtain the best homology Further adjustment was by computer graphic modelling of the altered yeast side chains on the tertiary structure of the rabbit muscle enzyme (Details of this procedure will be published elsewhere) Gaps (−) have been inserted to maximize similarity Only non-identical residues of rabbit muscle phosphorylase are indicated At the N terminus, yeast phosphorylase is 28 residues longer than the muscle enzyme, at the C terminus, it is shorter by 11 residues There are 12 other sites of insertions/deletions in the yeast sequence relative to muscle Two deletions, of 8 and of 3 residues, are found at positions corresponding to loops in the rabbit muscle enzyme (316–323, 723–725) Of the 10 insertions, 5 involve additions of a single amino acid, preceding positions 438, 474, 735, 767 and 794 of the rabbit enzyme sequence The rest are multiple-residue insertions three residues at position 507, six at 213 and 557, nine at 591 and 13 at 114 Lower case letters, residues of rabbit muscle phosphorylase involved in catalytic or regulatory interactions d, dimer or subunit interface, p, covalent phosphorylation site, a, AMP effector binding site, g, active site (and glucose inhibition), c, caffeine and purine inhibitor binding site, s, glycogen storage site, v, pyridoxal-phosphate cofactor binding site

Muscle glycogen phosphorylase exhibits allosteric activation by AMP, the binding site is well characterized in the three-dimensional structure and is 30 Å from the active site[15] The site can be thought of as three subsites, one each for the phosphoryl, sugar and base moieties of AMP The phosphoryl and sugar subsites also bind phosphate, glucose 1-phosphate and glucose 6-phosphate[16] The positive charges of Arg 242, Arg 309 and Arg 310, which form the pocket for binding the phosphate moiety of AMP in rabbit muscle phosphorylase, are conserved in the yeast sequence and, from model building, appear to form a similar binding pocket In the sugar subsite of muscle phosphorylase, Asp 42 and Gln 71 interact with the ribose portion of AMP, in the yeast enzyme, only Gln 71 is conserved In the base subsite, Tyr 75 forms an important stacking interaction with the adenyl group of AMP, and Asn 44 and Val 45 also make van der Waals' contacts with the adenine, none of these residues is conserved in the yeast enzyme It is unlikely that the activator site of muscle phosphorylase functions in the yeast enzyme by binding AMP, not only because important binding residues are not conserved, but because the protein structure near the sugar and base subsites is probably also altered All the residues mentioned which comprise these subsites lie within the N-terminal 80 amino acids of muscle phosphorylase, a region of little structural homology

A more intricate control of glycogen metabolism is exerted by reversible covalent phosphorylation of phosphorylases Although yeast and muscle phosphorylases interconvert between phosphorylated (active) and dephosphorylated (inactive) forms, many components of the mechanism are known to be different For example, the phosphorylated peptide isolated from the N terminus of rabbit muscle phosphorylase is conserved in vertebrate phosphorylases[17], the phosphorylated peptide in the yeast enzyme, however, does not match this sequence[18] In addition, the specific modifying enzymes of muscle phosphorylase do not act on the yeast enzyme[9] From the poor homology and different N terminus lengths (Fig 2), it is obvious that the substrate sites for the interconverting enzymes are dissimilar The target of phosphorylation in the yeast enzyme is a threonine (residue 19) that appears in the extended N terminus The amino-acid sequence around this residue (Arg-Arg-Leu-Thr) matches the consensus sequence for phosphorylation sites in substrates of cyclic AMP-dependent protein kinase[1], in accordance with the finding that yeast phosphorylase can be phosphorylated by this enzyme[19,20] We cannot identify any homologous element in the

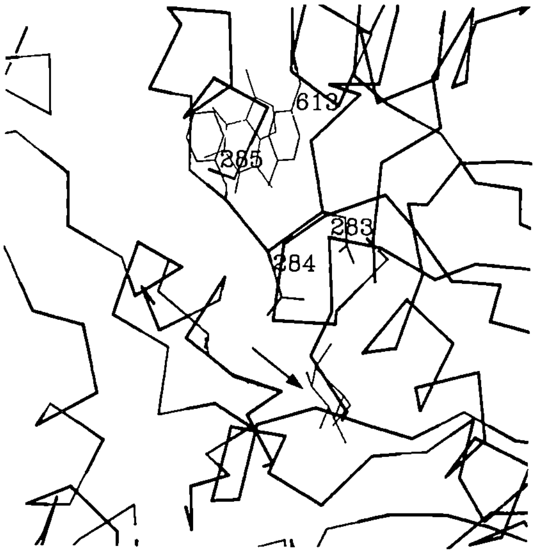


Fig 3 Comparison of the three-dimensional structures of rabbit muscle and yeast phosphorylases at the active-site cleft Darker segments of the α-carbon representation indicate identical or chemically similar residues conserved between the yeast and rabbit muscle enzymes, lighter segments indicate non-conserved residues Predominantly in dark segments, this region appears structurally conserved in the yeast enzyme Arrow, the active site, shown occupied by glucose and with conserved side chains Asp 283 and Asn 284 Side chains Phe 285 and Tyr 613 (thin lines) form the purine inhibition site (shown occupied by caffeine) and are also conserved in the yeast enzyme

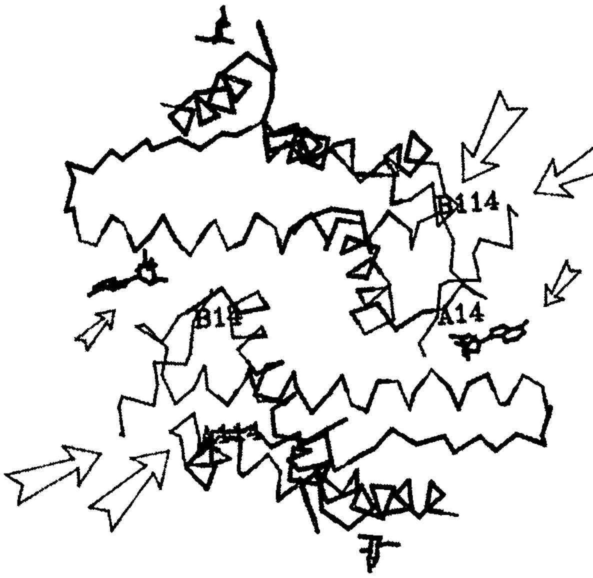


Fig. 4 Comparison of yeast phosphorylase to the three-dimensional structure of the rabbit muscle phosphorylase dimer (active conformation) at the AMP site, subunit interface and serine-14 phosphorylation site Shown are projected α-carbon representations of a similar region in two contacting subunits (chains distinguished by labelling residues A and B) Identical or similar residues between rabbit muscle and yeast enzymes are indicated in darker segments Lighter segments indicate non-conserved positions Many differences occur in this region The phosphorylation site, Ser 14 (positions A14 or B14), lies adjacent to the AMP site, shown occupied by AMP (small arrow) Larger arrows mark position 114, where yeast phosphorylase contains a 13-residue insertion, and the N terminus, where the yeast enzyme contains 28 extra residues

yeast sequence with which the phosphorylation site of muscle phosphorylase (Ser 14) might align In addition, there is no homology between the yeast N-terminal sequence and corresponding regions of other eukaryotic or prokaryotic phosphorylases$^{5,21-25}$

The intersubunit contacts of the rabbit muscle phosphorylase dimer are also significant in the regulatory mechanisms of the enzyme^{6,7} The structural segments that comprise the subunit interface couple the AMP and phosphorylation sites to the active-site region^{6,26} Major differences between muscle and yeast enzymes occur in this region (Fig 4) We have already noted specific alterations at the AMP binding site and the phosphorylation site, and the lack of homology at the N terminus Other residues involved in the intersubunit contact of muscle phosphorylase are indicated in Fig 2, most of these are not conserved in the yeast sequence In addition, there are two major length variations at positions near the dimer interface, at the N terminus, yeast phosphorylase has an extra 28 residues compared with the muscle homologue, and at a position homologous to residue 113 of the muscle enzyme, the yeast enzyme contains an insertion of 13 residues Thus, the subunit interface, as found in rabbit muscle phosphorylase, appears to be altered in the yeast enzyme, implying that intersubunit cooperativity in yeast phosphorylase involves structural components and mechanisms that are distinct from those in the muscle enzyme

Yeast phosphorylase is strongly inhibited by glucose 6-phosphate^{9,10} Although our analysis shows that a functional AMP activator site is disrupted in the yeast enzyme, conservation of the phosphoryl subsite and partial conservation of the sugar subsite suggest that this region serves as the binding site for glucose 6-phosphate This hypothesis is supported by the following evidence Inhibition of yeast phosphorylase by glucose 6-phosphate is non-competitive with respect to both glucose 1-phosphate and glycogen^{9}, implying that the effector binds somewhere other than at the catalytic site Glucose 6-phosphate does bind at the AMP site of inactive (dephosphorylated) muscle phosphorylase and inhibits activation (phosphorylation)^{16,27} Crystallographic studies^{16} have implicated all the conserved residues between yeast and muscle enzymes at the AMP site in the binding of glucose 6-phosphate to rabbit muscle phosphorylase Trp 67, another residue in contact with glucose 6-phosphate in rabbit muscle phosphorylase^{16}, is also conserved in the yeast enzyme

The lack of sequence homology at the N terminus indicates that the phosphorylation sites of yeast and vertebrate phosphorylases did not evolve from a common ancestral regulatory region Palm *et al*^{5} have suggested that a preformed nucleotide sequence encoding a phosphorylatable peptide was fused to a primordial phosphorylase gene in the evolution of vertebrate phosphorylases If true, it seems that a separate gene fusion event generated the phosphorylation site in yeast phosphorylase Of course, the acquisition of a phosphorylation site alone would not have been sufficient for covalent regulation, because allosteric regulation by reversible phosphorylation requires intersubunit interactions that couple phosphorylation to active-site changes The N-terminal sequence divergence that generated different phosphorylation sites in yeast and muscle enzymes also affected development of appropriate intersubunit contacts Although its structure remains to be elucidated, the subunit interface in yeast phosphorylase must differ significantly from that of muscle phosphorylase but still functions analogously, transmitting regulatory responses to the catalytic site Hence, the intersubunit interactions responsible for allosteric transmission in yeast and muscle phosphorylases seem to have arisen by convergent evolution Presumably, the dissimilar N termini also influenced the completion of the coupled allosteric ligand-binding site, enabling development of contrasting regulatory functions, glucose 6-phosphate inhibition in yeast and AMP activation in muscle

Finally, how has the regulatory phosphorylation cascade evolved with time? The phosphorylated and dephosphorylated forms of muscle phosphorylase are interconverted by phosphorylase kinase and phosphorylase phosphatase, specific enzymes that, in turn, are regulated (phosphorylated) by cyclic AMP-dependent protein kinase The phosphorylation site of yeast phosphorylase, however, is recognized by both the cyclic AMP-dependent protein kinase and a specific phosphorylase kinase[19,20], as well as by a general protein phosphatase[28,29] The following evolutionary model is consistent with these observations a phosphorylation site recognized by a general protein kinase (and phosphatase) was initially spliced onto the N terminus of the ancestral enzyme, divergence of this site led to poorer recognition and eventually loss of recognition by the general protein kinase, coevolution of other kinases (and phosphatases) gradually restricted their range of substrates, resulting in specificity for phosphorylase The difference in phosphorylation site specificity between yeast and muscle phosphorylases may reflect how far the regulatory cascades of yeast and vertebrates have evolved in this general path

We thank Drs Michael Stern, Susan Michaelis, Christopher Newgard, Kenichi Nakano and John Burke for helpful advice, discussions and technical assistance Supported by a grant from the NSF (PCM84-20489)

Received 19 May, accepted 7 August 1986

1 Cohen, P *Eur J Biochem.* **151,** 439-448 (1985)
2 Cohen, P *Control of Enzyme Activity* 2nd edn, 1-96 (Chapman & Hall, London, 1983)
3 Graves, D J & Wang, T H in *The Enzymes* Vol 7, 3rd edn (ed Boyer, P), 430-482 (Academic, London, 1982)
4 Fischer, E H , Pocker, A & Saari, J C *Essays Biochem.* **6,** 23-68 (1970)
5 Palm, D , Goerl, R. & Burger, K. J *Nature* **313,** 500-502 (1985)
6 Fletterick, R J & Sprang, S R *Acc. chem. Res* **15,** 361-369 (1983)
7 Fletterick, R J & Madsen, N B A *Rev Biochem* **49,** 831-861 (1980)
8 Dombradi, V *Int J Biochem.* **13,** 125-139 (1981)
9 Fosset, M , Muir, L W , Nielsen, L D & Fischer, E H *Biochemistry* **10,** 4105-4113 (1971)
10 Sagardia, F , Gotay, I & Rodriquez, M *Biochem biophys Res Comm.* **42,** 829-935 (1971)
11 Kasvinsky, P J , Schechosky, S & Fletterick, R. J *J biol. Chem.* **253,** 9102-9106 (1978)
12 Guenard, D , Morange, M & Buc, H *Eur J Biochem.* **76,** 447-452 (1977)
13 Stura, E A *et al. J molec. Biol.* **170,** 529-565 (1983)
14 Sprang, S *et al. Biochemistry* **21,** 2036-2048 (1982)
15 Sprang, S R & Fletterick, R. J *J molec. Biol.* **131,** 523-551 (1979)
16 Lorek, A *et al. Biochem. J* **218,** 45-60 (1984)
17 Titani, K *et al. Proc. natn. Acad Sci. U S A.* **74,** 4762-4766 (1977)
18 Lerch, K & Fischer, E H *Biochemistry* **14,** 2009-2014 (1975)
19 Becker, J -U , Wingender-Drissen, R. & Schiltz, E *Archs Biochem. Biophys* **225,** 667-678 (1983)
20 Pohlig, G , Wingender-Drissen, R. & Becker, J U *Biochem. biophys Res Commun.* **114,** 331-338 (1983)
21 Cohen, P , Saari, J C & Fischer, E H *Biochemistry* **12,** 5233-5241 (1973)
22 Schiltz, E , Palm, D & Klein, H W *FEBS Lett* **109,** 59-62 (1980)
23 Lacks, S A , Dunn, J J & Greenberg, B *Cell* **31,** 327-336 (1982)
24 Nakano, K., Fukui, T & Matsubara, H *J biol. Chem.* **255,** 9255-9261 (1980)
25 Newgard, C B , Nakano, K , Hwang, P K & Fletterick, R J *Proc. natn Acad Sci U S A.* (in the press)
26 Sprang, S & Fletterick, R J *Biophys. J* **32,** 175-192 (1980)
27 Madsen, N B & Scheckosky, S *J biol. Chem.* **242,** 3301-3307 (1967)
28 Wingender-Drissen, R & Becker, J -U *Biochim. biophys Acta* **743,** 343-350 (1983)
29 Matsumoto, K , Uno, I , Kato, K. & Ishikawa, T *Yeast* **1,** 25-38 (1985)
30 Carlson, M & Botstein, D *Cell* **28,** 145-154 (1982)
31 Southern, E *J molec. Biol.* **98,** 503-517 (1975)
32 Sanger, F , Coulson, A R , Barrell, B G , Smith, A J H & Roe, B A *J molec. Biol.* **143,** 161-178 (1980)
33 Nakano, K , Hwang, P K. & Fletterick, R. J *FEBS Lett* (in the press)
34 Dayhoff, M O (ed) in *Atlas of Protein Sequence and Structure* Vol 5, Suppl 2, 3-8 (Natn Biomed Res Fdn, Silver Spring, Maryland, 1976)

α-Lactalbumin possesses a novel calcium binding loop

D. I. Stuart*, K. R. Acharya*, N. P. C. Walker†‡, S. G. Smith*, M. Lewis*‡ & D. C. Phillips*

* Laboratory of Molecular Biophysics, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK
† Chemistry Department, Queen Mary College, Mile End Road, London E1 4NS, UK

Calcium performs a unique role in biology, achieving biological effects through highly specific interactions with and modulation of target proteins[1]. It has been proposed that calcium-modulated proteins possess a characteristic, evolutionarily related, binding fold, known as the EF-hand[2]. The high-resolution X-ray structure of α-lactalbumin reveals a Ca^{2+} binding fold that resembles an EF-hand only superficially and presumably has no evolutionary relationship with it. However, there is clear homology with the corresponding loop in c-type lysozyme (the 'parent' molecule of α-lactalbumin). This study, at 1.7 Å resolution, represents one of the most accurate analyses of a calcium binding protein yet reported.

α-Lactalbumin acts as a specifier protein[3], it comprises some 15% of total protein in human milk[4] and is essential for the production of lactose The attachment of α-lactalbumin to β-galactosyl transferase on the luminal surface of the Golgi membrane results in a complex, lactose synthetase that catalyses the addition of galactose to glucose to produce lactose α-Lactalbumin is a small protein that has about 40% sequence identity with c-type lysozyme, and also a similar intron-exon genetic organization[5], strongly suggesting that the two have diverged from a common ancestor α-Lactalbumin is a calcium metalloprotein[6], the K_{app} in physiological conditions is likely to be of the order of 10^{-6}–10^{-7} M (ref 7), comparable to that observed for the calcium-modulated proteins[1] In addition, Zn^{2+} binding has been observed (K_d 5 μM)[8] The biological role of metal ion binding is unclear If the concentration of free calcium ions in the Golgi lumen is comparable to that in milk (about 3 mM), it would be expected that the Ca^{2+} site would be fully occupied, whereas the occupation of the Zn^{2+} site would be uncertain The first role proposed for the Ca^{2+} binding was to stabilize the protein and protect it from thermal denaturation[6,9], more recently it has been claimed that the pattern of ion binding affects the catalytic properties (notably the V_{max}) of the lactose synthetase complex[8] Calcium binding may also be important in controlling the release of α-lactalbumin from the Golgi membrane, an essential step in inducing lactation[8] It is plausible that these mechanisms allow fine control of lactose synthesis

The second messenger role of Ca^{2+} ions, binding to target proteins and hence inducing a conformational change, is well established Unfortunately, in no case is the precise mechanism of triggering understood in terms of the concomitant conformational changes, although the crystal structure of troponin-C gives some hints[10] In the case of α-lactalbumin, knowing the structure of lysozyme, a close homologue with no marked propensity for binding calcium, it is possible to infer something of the conformational changes induced by the calcium binding by comparison with the structure of the calcium-bound form of α-lactalbumin

Structural studies on crystalline α-lactalbumins have been in progress for a number of years but have been frustrated by various problems[11,12] We have determined the structure of α-lactalbumin from baboon milk Low-resolution isomorphous replacement phases allowed the approximate orientation of the molecule in the unit cell to be determined and confirmed that the structure is broadly similar to lysozyme[13,14] A process of careful refinement (initially of a lysozyme model) against the X-ray observations revealed the detailed structure This work will be reported in detail elsewhere The sequence of the enzyme from baboon milk has not been determined, but preliminary experiments suggested very few differences between the sequences of human and baboon α-lactalbumins (R Greenberg, personal communication) Our present model therefore has few changes from the human α-lactalbumin sequence The *R*-factor[32] of 0 207 (for 10,627 data $>2\sigma(F)$ to a resolution of 1 7 Å) together with the reasonable stereochemistry exhibited by the model (r m s (root mean square) deviation from ideal

---

‡ Present addresses BASF AG, Hauptlaboratorium A30, D 6700 Ludwigshafen, FRG (N P C W), Smith Kline and French Laboratories, 1500 Spring Garden Street, P O Box 7929, Philadelphia, Pennsylvania 19101, USA (M L)

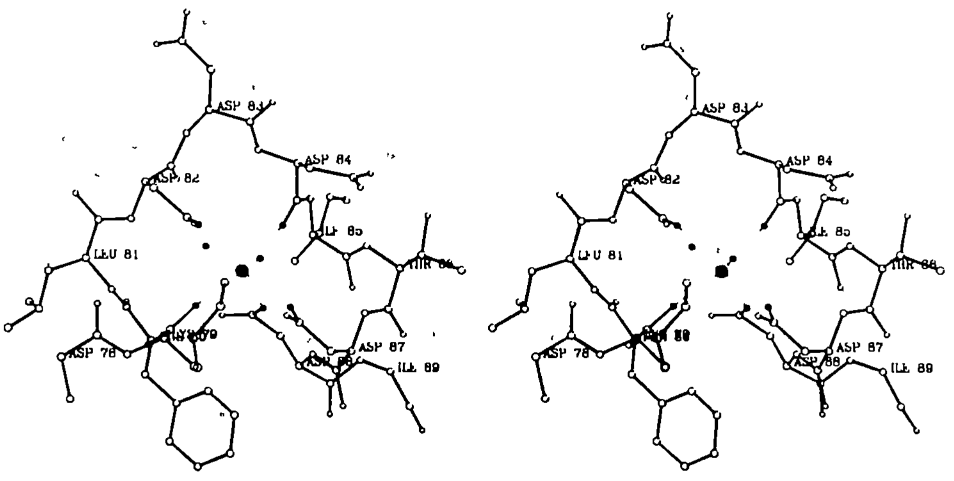


Fig. 1 Stereoscopic views of the calcium binding loop in α-lactalbumin. Residues 78–89 plus the calcium ion (larger black dot) and its water ligands. Dashed lines join the calcium to its ligands.

covalent bond lengths = 0.03 Å) and the high quality of the electron density map lead us to be confident of our interpretation of the molecular structure. We have placed all the protein atoms and 132 water molecules. We estimate that the coordinates for the parts of the structure that we discuss below are accurate to within a few tenths of an ångstrom unit. The determination of the α-lactalbumin structure represents a notable application of the molecular replacement method as the starting model has only 37% sequence identity with α-lactalbumin and is a distinct enzyme.

When the enzyme is crystallized and no metals added to the crystallization medium, as expected from the binding affinity data, the calcium binding site is occupied. Nucleic acid sequences[15] and revised amino-acid sequences[16] reveal two pairs of strongly conserved aspartic acid residues (82, 83 and 87, 88). This led to the proposal[16], subsequently supported by NMR data[17], that these four residues are concerned in calcium binding. We find that three of these residues are calcium ligands. The Ca^{2+} site is easily identifiable as a persistent feature of our difference Fourier maps which cannot be adequately accounted for by a water molecule. Refinement fully confirms this assignment, the current B-values for the ion and its ligands being about 15 Å^2 and indeed the B-factors for all this region are among the lowest in the molecule. Figure 1 shows the detailed atomic structure around the calcium, which is surrounded by seven ligands, all oxygens, three from the carboxylate groups of aspartyl residues (82, 87 and 88), two carbonyl oxygen atoms (from residues 79 and 84) and two water molecules. The five ligands contributed by the protein come from a rather tight turn akin to a β-turn with an extra residue, Asp 82, acting as a 'spacer' between two helices. We shall subsequently refer to this structure as the 'elbow'. The geometry in small-molecule-Ca^{2+} complexes is variable but generally involves six, seven or eight ligands with a distorted pentagonal bipyramid being the most common configuration in the seven coordinate case (G. Klebe, personal communication). When calcium binds to proteins, the solvent often contributes one or more ligand[18]. In α-lactalbumin the coordination is ideal, forming a slightly distorted pentagonal bipyramid with carbonyl oxygen atoms at the apices while the carboxylate groups and water are arranged with the waters approximately opposed (Fig. 1). The Ca–O distances, ranging from 2.3 Å to 2.5 Å (Table 1), are close to the expected value of 2.4 Å.

The only comparable calcium binding sites described for other proteins (in terms of affinity and topology) are those forming the class of folds known as the EF-hand[2]. Comparison of the α-lactalbumin loop with an EF-hand (see Fig. 2) reveals that although the elbow of α-lactalbumin is a more compact structure (consisting of two fewer residues) there are certain marked similarities, both structures are formed from a loop between almost orthogonal helices and deploy similar types of ligands around the Ca^{2+}. In α-lactalbumin a single Asx turn[19] stabilizes the structure whilst the longer EF-hand uses several of these. Furthermore, in both cases there is a conserved central hydrophobic residue (Ile 85 in α-lactalbumin) which is important in stabilizing the structure as a whole. We can discount the naive hypothesis of an homologous relationship for several reasons. Firstly because the evolutionary history of α-lactalbumin rules out such a scheme. Secondly, the Ca^{2+} coordination observed for the typical EF-hand is considered a distorted octahedron compared with the pentagonal bipyramid of α-lactalbumin. Finally, on searching a database of well-refined structures for related conformations[20], the only structures found that were closely similar to a given EF-hand were other EF-hands. These structures superimpose to better than 1 Å r.m.s. deviation. As there are a large number of unrelated structures that superimpose to within about 2 Å, there is no evidence for significance in the structural similarity with the elbow where the r.m.s. deviation exceeds 2 Å.

Comparing the calcium binding loop in α-lactalbumin with the corresponding structure in hen egg white lysozyme we find, surprisingly, that the conformation of the backbone is essentially unchanged whereas the side chains are radically altered (Fig. 2). The r.m.s. deviation in C_α positions for this loop is 1.2 Å. The differences are largely attributable to movement of the loop away from the body of the molecule in α-lactalbumin compared with lysozyme. If we superimpose the coordinates of the loop alone, we find that they agree with an r.m.s. deviation of 0.6 Å, as close a similarity as that found between the calcium binding loops of the EF-hand proteins[18] and indeed as that between human and hen lysozyme. Further evidence of the high degree of similarity comes from an analysis of the water structure around this loop in α-lactalbumin and lysozyme. There is a network of three water molecules in hen egg white lysozyme that corresponds to two water molecules and the calcium ion in α-lactalbumin[21]. The fold at the elbow seems to be unusual, no other loops superimpose well. The best fit (with rhodanese, residue 278 onwards) has an r.m.s. deviation of 1.8 Å and many others are about as good. Hence, although this loop is a region of considerable similarity in structure, there are important differences in sequence (Table 2).

Table 1 The calcium ligands

Residue		Group		Distance (Å)
79	Lys	Carbonyl	O	2.3
82	Asp	Carboxylate	OD1	2.4
84	Asp	Carbonyl	O	2.3
87	Asp	Carboxylate	OD1	2.3
88	Asp	Carboxylate	OD1	2.4
		Water	O	2.4
		Water	O	2.5

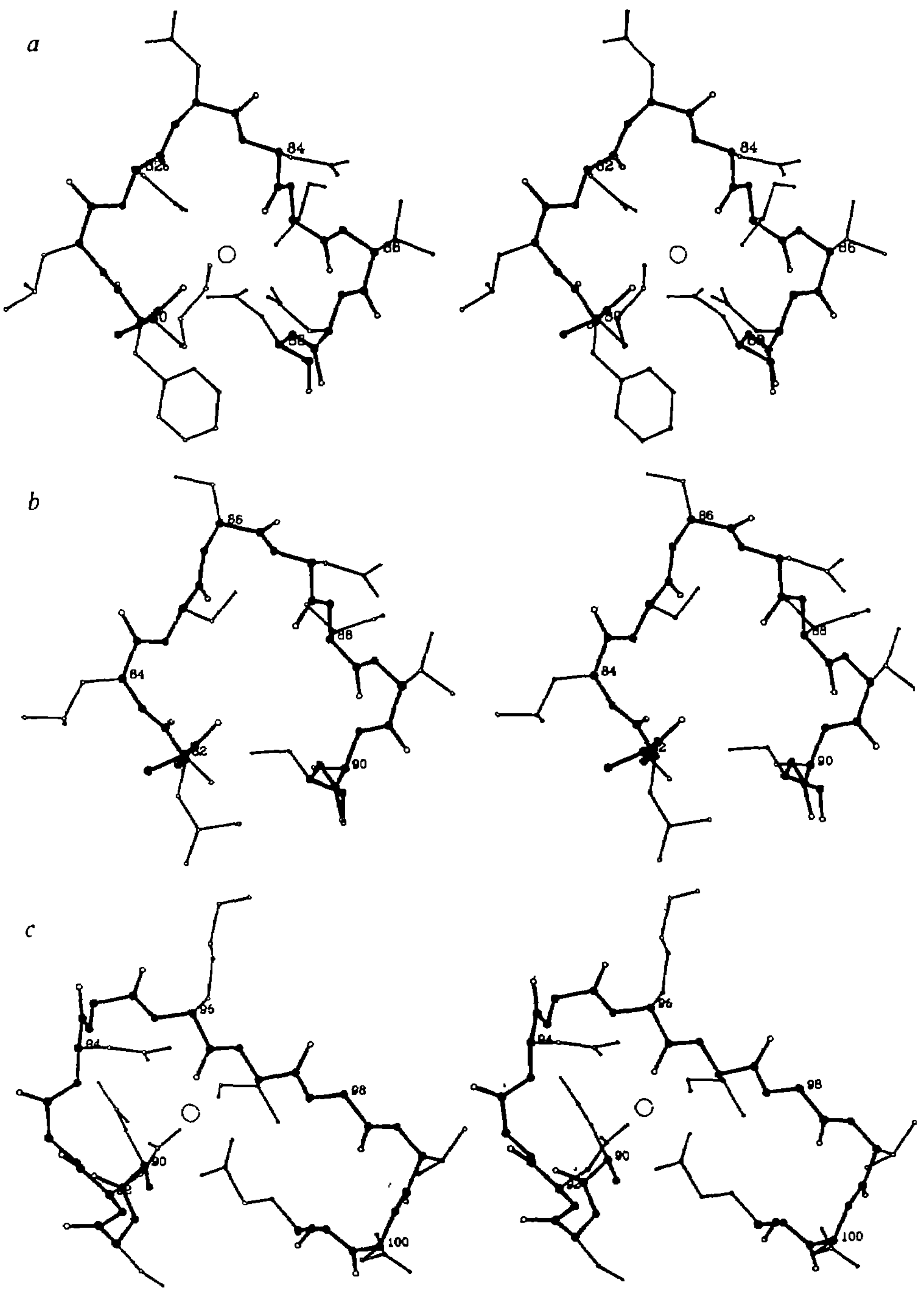


Fig. 2 Rotationally aligned structures of *a*, residues 79-88 in α-lactalbumin, *b*, residues 82-91 in hen lysozyme, *c*, residues 90-101 in carp parvalbumin, the EF-hand, data set F6A in the protein data bank[29] The central, enveloped atom in *a* and *c* is calcium The superimpositions were performed using a computer program[30] drawing on the ideas of Rossmann and Argos[31] The optimal rotation matrices were determined on the basis of the C_α coordinates of the residues in the loops (all 10 residues of the α-lactalbumin loop were matched with parvalbumin residues to give an r m s deviation of 2 5 Å)

The above analysis refers to the hen lysozyme structure We have also looked at human lysozyme[22] where the three aspartates of α-lactalbumin that contribute carboxylate ligands correspond to Gln 86, Asp 91 and Ala 92 Asp 91 is rotated by ~90° about $\chi 1$ (compared to α-lactalbumin), to bring the charged atoms into solution, whereas the ε-nitrogen of Gln 86 is at the position of a water ligand of Ca^{2+} in α-lactalbumin It seems therefore that human lysozyme contains a partially formed calcium binding site without a destabilization of protein structure Ala 92 is rather internal and its conversion to an Asp in the absence of some compensating charge or change in conformation would be expected to destabilize the protein A survey of lysozyme sequences reveals that the sequence of horse milk lysozyme[23] shows the full complement of the three aspartate residues that form Ca^{2+} ligands in α-lactalbumin If this sequence is correct, it raises the intriguing possibility that horse milk lysozyme might bind calcium

It has been suggested that the entropic effects of calcium binding are more unfavourable in α-lactalbumin than the EF-hand proteins[24] At first sight this is puzzling, the *B*-values for this loop in lysozyme are low[21], and therefore entropic effects due to ordering of the loop on calcium binding should be small Furthermore the rigidity of the calcium binding loop in α-lactalbumin is responsible for the discrimination the site shows between Ca^{2+} and Mg^{2+} (the K_ds differ by approximately 10^5), the loop being unable to collapse around the smaller Mg^{2+} ion Nevertheless, uncompensated negative charges should destabilize the observed lysozyme conformation, so what conformational changes are likely? The results of the database search do not indicate any closely related folds that could be formed by rearrangement of the loop but a variety of structures are observed that differ by about 2 Å No quantitative evidence is available on the Ca^{2+}-induced conformational changes[25] but a possible mode of action could involve a relatively small conformational change at the Ca^{2+} binding site, perhaps triggering rigid body movements of secondary structural units as has been described for other structural transitions[26]

The observations on the structure of the calcium binding

Table 2 Amino-acid sequences at the α-lactalbumin elbow

Protein	79	80	81	82	83	84	85	86	87	88
α-lac, human	Lys	Phe	Leu	Asp	Asp	Asp	Ile	Thr	Asp	Asp
α-lac, bovine	Lys	Phe	Leu	Asp	Asp	Asp	Leu	Thr	Asp	Asp
Lys, hen	Ala	Leu	Leu	Ser	Ser	Asp	Ile	Thr	Ala	Ser
Lys, human	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Ala
Lys, horse	Lys	Leu	Leu	Asp	Glu	Asn	Ile	Asp	Asp	Asp
Calcium ligands in α-lac	+			*		+			*	*

The residue numbering is based on human α-lactalbumin Add three to these numbers to get the classical hen lysozyme numbering
* Residues that contribute a side-chain calcium ligand, + residues that contribute a carbonyl oxygen ligand

elbow are relevant to protein evolution The divergence of α-lactalbumin from lysozyme involves the formation of a novel enzyme system, and seems to be related to the formation of an extremely well developed Ca^{2+} binding ability, a surprising innovation We would argue against an alternative evolutionary interpretation, that an ancestral lysozyme possessed Ca^{2+} binding activity and the current lysozyme structure retains a 'memory' of this, because the proposed serial homologues (phage, goose and hen type lysozymes[27]) do not possess this loop as a common feature It is an interesting coincidence, however, that the loop of phage lysozyme once proposed as a calcium binding loop is probably homologous to a region fairly close to the elbow in α-lactalbumin[27,28] Clearly the evolution of a novel functional property of the α-lactalbumin molecule has not come about through exon shuffling or any sophisticated structural modifications It seems instead to be a structurally localized property that arose principally through simple amino-acid substitutions at three positions to produce a calcium binding site of exquisite specificity While other cases of divergence of structures to fulfil radically different functions are known (for example, the various functionally diverse immunoglobulin domains), here the process is so recent that we feel that we have almost caught evolution *flagrante delicto* Naturally α-lactalbumin may lend support to the rather simplistic approach to protein engineering that considers function to be a local property Such a hypothesis can be tested easily in this case by investigating the effects of mutagenic modification of the lysozyme loop in an attempt to produce a calcium binding site (as may have already been accomplished naturally in the horse milk lysozyme) We would be pleased if our observations stimulate such studies and studies on the biological significance of calcium binding to the lactose synthetase Although the lactose synthetase complex is not cytosolic, it is interesting to see essentially intracellular calcium binding at a site with no homology to the EF-hand At any event it is encouraging to see nature innovating rather than restricting itself to spare-functional-part protein design

This work was supported by the MRC and SERC

Received 14 April, accepted 18 August 1986

1 Levine, B A & Williams, R. J P in *Calcium and Cell Function*, Vol II (ed Cheung, W Y) 1-38 (Academic, New York, 1982)
2 Kretsinger, R. H & Nockolds, C E *J biol Chem.* **248**, 3313-3326 (1973)
3 Brew, K & Campbell, P N *Biochem. J* **102**, 258-269 (1967)
4 Aschaffenburg, R & Drewry, J *Biochem. J* **65**, 273-277 (1957)
5 Qasba, P K & Safaya, S K *Nature* **308**, 377-380 (1984)
6 Hiraoka, Y *et al Biochem. biophys Res Commun.* **95**, 1098-1104 (1980)
7 Mitani, M *et al J biol. Chem* **261**, 8824-8829 (1986)
8 Musci, G & Berliner, L J *Biochemistry* **24**, 6945-6948 (1985)
9 Hiraoka, Y & Sugai, S *Int J Pept Prot Res* **23**, 535-542 (1984)
10 Herzberg, O , Moult, J & James, M N G *J biol Chem.* **261**, 2638-2644 (1986)
11 Aschaffenburg, R *et al. J molec. Biol.* **67**, 525-528 (1972)
12 Fenna, R E *J molec. Biol* **161**, 203-210 (1982)
13 Aschaffenburg, R *et al. J molec. Biol.* **127**, 135-137 (1979)
14 Smith, S G *et al. Biochem. J* (submitted)
15 Hall, L *et al. Nucleic Acid Res* **10**, 3503-3515 (1982)
16 Shewale, J G , Sinha, S K. & Bew, K. *J biol. Chem.* **259**, 4947-4956 (1984)
17 Kuwajima, K , Hiroshima, Y & Sugai, S *Int J Pept Prot Res* **27**, 18-27 (1986)
18 Herzberg, O & James, M N G *Biochemistry* **24**, 5298-5302 (1985)
19 Rees, D C , Lewis, M & Lipscomb, W N *J molec. Biol.* **168**, 367-387 (1983)
20 Jones, T A & Thirup, S *EMBO J* **5**, 819-822 (1986)
21 Handoll, H D *thesis*, Univ Oxford (1985)
22 Artymiuk, P J & Blake, C C F *J molec. Biol.* **152**, 737-762 (1981)
23 McKenzie, H A & Shaw, D C *Biochem. Int* 10, 23-31 (1985)
24 Schaer, J-J, Milos, M & Cox, J A *FEBS Lett* **190**, 77-80 (1985)
25 Koga, K & Berliner, L J *Biochemistry* **24**, 7257-7262 (1985)
26 Chothia, C & Lesk, A M *Trends biochem. Sci.* **111**, 116-118 (1985)
27 Weaver, L H *et al. J molec. Evol.* **21**, 97-111 (1985)
28 Tufty, R M & Kretsinger, R H *Science* **187**, 167-169 (1975)
29 Bernstein, F C *et al. J molec. Biol.* **112**, 535-542 (1977)
30 Stuart, D I thesis, Bristol Univ (1979)
31 Rossmann, M G & Argos, P *J molec. Biol.* **109**, 99-129 (1977)
32 Blundell, T L & Johnson, L N in *Protein Crystallography*, 147 (Academic, New York, 1976)

Bent DNA at a yeast autonomously replicating sequence

Michael Snyder*†, Andrew R. Buchman‡ & Ronald W. Davis†

* Department of Biology, Yale University, New Haven, Connecticut 06511, USA
† Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA
‡ Department of Cell Biology, Stanford University, Stanford, California 94305, USA

DNA fragments that show retarded electrophoretic mobility through polyacrylamide gels have been found in both prokaryotes and eukaryotes[1-7]. In the case of kinetoplast DNA, evidence has been presented that the DNA is curved or 'bent'[8,9]. Bent DNA has previously been found at the λ and simian virus 40 (SV40) DNA replication origins[6,7]. Here we show the existence of bent DNA at a yeast autonomously replicating sequence (ARS1), a putative replication origin. The bent DNA has been localized to a 40–55 base pair (bp) segment and contains six $(A)_{3-5}$ stretches (that is, six poly(A) stretches, three to five nucleotides in length) phased approximately every 10.5 bp. This region contains a DNA binding site for a yeast protein factor. This site lies at the 3′ end of the *TRP1* gene, in a region devoid of nucleosomes, and is positioned 80 bp away from the ARS consensus sequence; removal of this region impairs ARS function *in vivo*. The bent DNA may be involved in transcription termination or the prevention of nucleosome assembly in this region.

Yeast ARS sequences have properties expected of origins of replication they allow the autonomous replication of sequences to which they are attached, occur at a frequency expected of chromosomal origins, replicate once per cell cycle using genes responsible for chromosomal replication and act as origins of replication *in vitro* (see ref 10 for review) Electron microscopic data for ribosomal DNA and 2μ DNA has mapped nascent replication bubbles near ARS elements suggesting that ARSs act as replication origins *in vivo*[11,12]

The structure of ARS1, which is near the 3′ end of the yeast *TRP1* gene, is shown in Fig 1 There are three elements that comprise ARS1, domains A, B and C Domain A is an 11-bp consensus sequence found at most other ARS sequences[13,14] Mutations in this sequence at ARS1 or the ARS at the HO locus obliterate ARS function *in vivo*[13,15,16] Domains B and C flank ARS1 and are important for its function[16-18], mutations in either

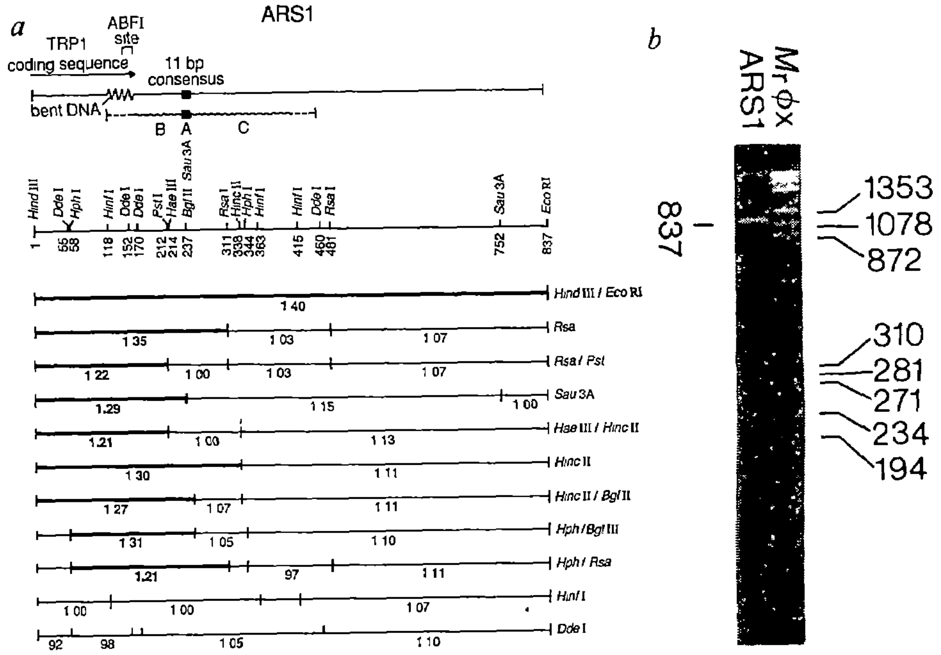


Fig. 1 *a*, Features of yeast ARS1; *b*, mobility of ARS1 restriction fragments on a 10% polyacrylamide gel. A (box), B and C (〰 lines) are indicated. The broken lines indicate the degree of uncertainty of the extent of domains B and C. The *Hin*dIII/*Eco*RI ARS1 fragment was digested with the various restriction enzymes indicated, ethanol-precipitated and the electrophoretic mobilities of the fragments measured in a 10% polyacrylamide gel using 50 mM Tris base, 50 mM boric acid, 1 mM EDTA buffer[20]. Electrophoresis was at 2.9 V cm^{-1} for 16 h at 24.5 °C. Correct assignment of the DNA bands was determined by digestion with a second enzyme. Beneath the fragments is indicated the *k* value = expected distance migrated/observed migration distance. Fragments where $k > 1.20$ are indicated in bold type (not determined for fragments less than 60 bp). *b*, Migration of the 837 bp (determined from DNA sequence[23]) *Hin*dIII/*Eco*RI ARS1 restriction fragment in a 10% polyacrylamide gel. 837 indicates the position of the 837-bp ARS1 fragment. Sizes of ΦX174 *Hae*III marker fragments are given. The 281-bp fragment of ΦX174 migrates more slowly than expected. Other standards used are λ *Bst*EII markers and pBR322 *Hin*fI markers (not shown).

region reduce, but do not abolish, ARS function *in vivo*. Domain B begins very close to the 3′ end of the *TRP1* gene and may overlap it. A protein factor called ABFI (ARS binding factor I) that has been partially purified from yeast extracts,[20] binds in this region. The roles of the different ARS domains A, B and C, and of ABFI are unknown.

Because bent DNA has been found at bacteriophage and viral DNA replication origins[6,7], we tested for unusual DNA conformational features in the ARS1 region by subjecting restriction fragments of ARS1 DNA to electrophoresis in 10% polyacrylamide gels and identifying DNA fragments that exhibited unusually retarded mobility[19,21,22]. The mobilities were compared to 'normal' DNA fragments derived from pBR322, ΦX174 and λ DNAs (see Fig. 1 legend). Several DNA fragments were found that migrate 20–40% more slowly than predicted from the known DNA sequence[23] at 24.5 °C (Fig. 1). For example, an 837-bp fragment migrates to a position expected of a 1,170-bp fragment, 40% more slowly than anticipated. Using the method of Diekmann and Wang[19], several restriction fragments spanning the ARS1 region were tested for abnormal electrophoretic mobility and their *k* values determined (*k* = expected distance migrated/observed distance migrated for the DNA fragments). All DNA fragments that contain the 150-bp region between the *Hph*I and *Pst*I sites are found to have *k* values > 1.20. DNA fragments generated by cleaving the DNA with *Hin*fI and *Dde*I that cut within this 150-bp segment do not show abnormal mobility. Data presented below demonstrate that the *Hin*fI (118) and *Dde*I (152, 170) sites bracket a region which has properties characteristic of bent DNA. It is clear that there is a strong correlation between the position of the 118–170 *Hin*fI/*Dde*I region within a fragment and the observed *k* value. It has been noted that when the bent sequence is near the end of a fragment, the effect on retarded mobility is diminished[21]. Similar results were obtained after the digested ARS1 DNA was extracted with phenol or treated with proteinase K (data not shown). Thus reduced mobility does not appear to be due to association with proteins.

The abnormal DNA region was further localized by deletion analysis. A series of five deletions beginning at the *Hin*dIII site and extending increasing distances towards ARS1 were constructed (Fig. 2). These are called Δ108–Δ197, the numbers refer to the number of base pairs deleted. Restriction fragments of plasmids containing these deletions were generated by cleaving at sites located 150–200 bp outside the deletion regions in such a way that the putative bend is approximately in the centre of the fragment. The mobility and *k* values of these DNA fragments are shown in Fig. 2. Both the wild-type (wt) and Δ108 DNA fragments migrate more slowly than expected (31 and 32% more slowly, respectively). Δ149 (which differs from Δ108 by 41 bp) migrates normally. Thus at least some of the DNA that confers abnormal mobility lies between 108–149 bp on the map in Fig. 1.

Bent DNA tends to migrate more 'normally' at higher temperatures[19,22]. ARS1 DNA fragments from the deletion plasmids were analysed as above except that the electrophoresis was performed at 38 °C instead of 25 °C. As shown in Fig. 2, the wild-type and Δ108 fragments migrated faster at the higher temperature (lower *k* values). Relative mobilities of fragments lacking the 108–149 region were not affected by temperature. Thus, the effect of temperature on abnormal migration is consistent with the presence of bent DNA. Abnormal migration cannot be due to simple denaturation of DNA sequences, which would have the opposite effect with temperature. Inspection of the DNA sequence in the 108–149 region reveals the presence of six A_{3-5} stretches phased approximately every 10.5 bp (Fig. 3). This feature is characteristic of the bent DNA observed in kinetoplast DNA and λ replication origin[6,8,21]. It has been shown[22] that bending occurs at the 3′ junctions of A tracts, and is maximized when the 3′ A junctions are phased approximately every 10 bp.

Interestingly, DNA footprint analysis with partially purified yeast extracts demonstrates that a protein factor (ABFI) specifically binds to this region at residues 139–155 (Fig. 3). So far, five different binding sites for this factor have been identified in yeast DNA near ARS elements[24]. The binding site at ARS1 contains an $(A)_5$ stretch directly in its centre and other binding sites contain $(A)_{3-4}$ stretches. The bent conformation of the DNA at ARS1 may affect the affinity with which this factor binds DNA, similarly to SV40 T antigen where the nucleotides contacted by T antigen flank a DNA bend that is critical for binding[7].

To determine whether the bent DNA region is important for ARS1 function *in vivo*, the effect of the five deletions shown in Fig. 2 on ARS activity was assayed. Each ARS fragment was

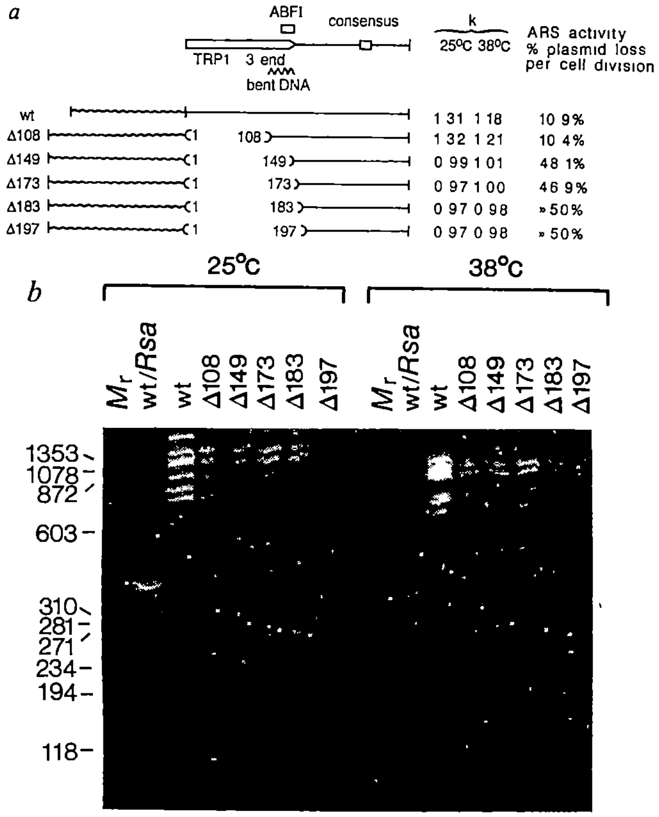

a
ABFI
consensus
TRP1 3 end
bent DNA
k
25°C 38°C
ARS activity % plasmid loss per cell division
wt 1 31 1 18 10 9%
Δ108 1 108 1 32 1 21 10 4%
Δ149 1 149 0 99 1 01 48 1%
Δ173 1 173 0 97 1 00 46 9%
Δ183 1 183 0 97 0 98 »50%
Δ197 1 197 0 97 0 98 »50%
b
25°C
38°C
M_r wt/Rsa wt Δ108 Δ149 Δ173 Δ183 Δ197
M_r wt/Rsa wt Δ108 Δ149 Δ173 Δ183 Δ197
1353
1078
872
603
310
281
271
234
194
118

Fig 2 Deletion analysis through the bent DNA region *a*, The features of ARS1 and deletion fragments, *b*, mobility of digested deletion plasmids *a*, All fragments were cloned into YRP14/CEN4[25] The wild-type (wt) fragment was inserted between the *Hin*dIII and *Eco*RI sites Deletions begin at the *Hin*dIII site and extend to the nucleotides indicated *Bam*HI octanucleotide linkers were added and the *Bam*HI/*Eco*RI fragments cloned into YRP14/CEN4 For bent DNA analysis wt DNA was cleaved with *Rsa*I and *Eco*RV which leaves 156 bp of pBR322 sequences (〰) adjacent to the *Hin*dIII/*Rsa*I ARS1 fragment For the deletion plasmids, DNA was digested with *Rsa*I and *Sph*I which leave 187 bp of pBR322 sequences adjacent to ARS1 fragments The lengths of the resulting fragments of interest are wt 467 bp, Δ108, 396 bp, Δ149, 355 bp, Δ173, 331 bp, Δ183, 321 bp, Δ197, 307 bp Electrophoresis was as in Fig 1, at 25 °C DNAs were also digested with enzymes that remove the pBR322 sequences and analysed on the same gel The wt fragment still migrates abnormally, but all deletion fragments Δ108-Δ197 now migrate with *k* values <1 05 (not shown) ARS activity for each plasmid was assayed by transforming the plasmids into yeast strain YNN318 (a/α ura3-52/ura3-52 lys2-801/lys2-801 ade2-101/ade2-101 trp1-Δ901/trp1-Δ901 his3-Δ200/his3-Δ200) Cells harbouring plasmids were transferred from selective yeast minimal plates containing glucose to yeast minimal media supplemented with galactose (Calbiochem) and grown for 16 h at 30 °C (selecting for the plasmid) and an additional 10 h without selection, then plated on nonselective galactose plates containing limiting adenine, cells containing the plasmids appear pink while those lacking the plasmid are red[25] The frequency of $\frac{1}{2}$ sectored colonies among colonies containing a single plasmid ($n=296$ to 2,000) was measured to determine the rate of plasmid loss per cell division Loss for wt and Δ108 plasmids was 4/1 1 0/2 0 segregation events while for Δ149 and Δ173 loss was >45/1 1 0/2 0 segregation events Loss rates are very high for Δ183 and Δ197 constructs, all colonies appear homogeneously red For Δ183 and Δ197 the number of plasmid-containing colonies plated was determined using selective glucose-containing plates When similar ARS activity experiments were performed using glucose media, the relative stabilities of the wt and Δ108-Δ197 plamids are unchanged, however, the overall plasmid stability is increased This medium effect is consistent with our observations that ARS plasmids are more stable on richer carbon sources[25,32] *b*, Polyacrylamide gel showing mobility of the digested deletion plasmids Left, DNAs analysed at 25 °C for 16 h, right, identical gel electrophoresis at 38 °C for 13 h Lanes are as indicated and described above Additional lanes are M_r, ΦX174 *Hae*III digested DNA (sizes indicated in bp), wt/*Rsa*, *Rsa*I the *Hin*dIII/*Eco*RI ARS1 DNA digested with *Rsa*I The 311-bp DNA fragment migrates coincident with a 356-bp DNA fragment at 25 °C (the 356-bp DNA fragment migrated 11% more slowly than expected on this particular gel)

cloned into YRP14/CEN4, a centromere-containing plasmid lacking an efficient replicator, and the rate of plasmid loss was determined by a colony colour assay (Fig 2)[25] Plasmids containing bent DNA (wild type and Δ108) are stable and lost at the rate of 10% per cell division The Δ149 plasmid, which lacks most of the bent DNA and half of the ABFI binding site, is fivefold less stable (Fig 2) Thus, the bent DNA region is important for ARS function *in vivo* It remains to be determined whether this effect is due to the bent DNA sequence, to the binding site for ABFI, or to both

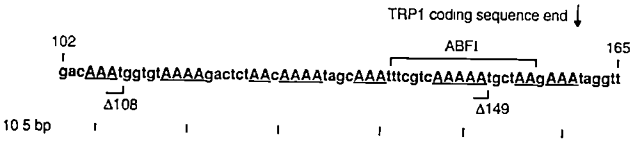

TRP1 coding sequence end
102
ABFI
165
gacAAAtggtgtAAAAgactctAAcAAAAtagcAAAtttcgtcAAAAAtgctAAgAAAtaggtt
Δ108
Δ149
10 5 bp

Fig 3 DNA sequence in the bent DNA region of ARS1 All A residues that have neighbouring A residues are shown in bold type 10 5-nucleotide intervals and the ABFI binding site are also indicated A 25-bp region centred around this site is protected in DNase I protection experiments[20] The endpoints of deletions Δ108 and Δ149 are shown The ARS1 consensus sequence is located at position 241-251

Bent DNA has been previously discovered at the bacteriophage λ and SV40 replication origins[6,7,26] The discovery of such DNA at yeast ARS1 suggests that it may be important at chromosomal origins as well Although the precise role of the bent DNA at ARS1 is presently unknown, the position of this region suggests several interesting possibilities Since it is located at the 3′ end of the *TRP1* gene, it may be important for transcription termination Transcription through ARS1 has been found to reduce its activity (M S and R W D, unpublished observations) Thus one role of domains B and C might be to act as transcription termination segments to protect domain A The bent DNA with its associated protein could serve an important role in this process In addition, the region immediately adjacent to the ARS1 consensus element on a *TRP1* ARS1 plasmid (nucleotides 60-242) has been shown to be free of nucleosomes[27,28] Another function of the DNA bend might be to keep this region free of nucleosomes and readily accessible to the DNA replication apparatus and/or other proteins The origin region of SV40 DNA has also been found to be free of nucleosomes[29-31]

We thank Ron Sapolsky for his help with sequence analysis and Roger Kornberg in whose laboratory some of these experiments were performed This research was supported by NIH grant GM 21891

Received 6 May, accepted 2 September 1986

1 Bossi, L & Smith, D M *Cell* **39**, 643-652 (1984)
2 Simpson, L *Proc natn Acad Sci USA* **76**, 1585-1588 (1979)
3 Ross, W, Schulman, M & Landy, A *J molec Biol* **156**, 505-529 (1982)
4 Stellwagen, N C *Biochemistry* **22**, 6186-6193 (1983)
5 Chalberg, S S & Englund, P T *J molec Biol* **138**, 447-472 (1980)
6 Zahn, K & Blattner, F R *Nature* **317**, 451-453 (1985)
7 Ryder, K, Silver, S, DeLucia, A L, Fanning, E & Tegtmeyer, P *Cell* **44**, 719-725 (1986)
8 Marini, J C, Levene, S D, Crothers, D M & Englund, P T *Proc natn Acad Sci USA* **79**, 7664-7668 (1982)
9 Hagerman, P J *Proc natn Acad Sci USA* **81**, 4632-4636 (1984)
10 Campbell, J L *Genet Engng* **5**, 109-155 (1983)
11 Newlon, C S, Devenish, R J, Suci, P A & Roffis, C J *ICN-UCLA Symp molec cell Biol* **22**, 501-516 (1981)
12 Saffer, L D & Miller, O L *Molec cell Biol* **6**, 1148-1157 (1986)
13 Stinchcomb, D C, Mann, C, Selker, E & Davis, R W in *The Initiation of DNA Replication and Segregation of Yeast Chromosomes* Vol 3 (ed Ray, D S) 473-488 (Academic, New York, 1981)
14 Broach, J R *et al Cold Spring Harb Symp Quant Biol* **47**, 1165-1173 (1982)
15 Celniker, S E, Sweder, K., Srienc, F, Bailey, J E & Campbell, J L *Molec cell Biol* **4**, 2455-2466 (1984)
16 Kearsey, S *Cell* **37**, 299-307 (1984)
17 Koshland, D, Kent, J C & Hartwell, L H *Cell* **40**, 393-403 (1985)
18 Srienc, F, Bailey, J E & Campbell, J L *Molec Cell Biol* **5**, 1676-1684 (1985)
19 Diekmann, S & Wang, J C *J Molec Biol* **186**, 1-11 (1985)
20 Buchman, A, Kimmerly, W, Rine, J & Kornberg, R, (manuscript in preparation)
21 Wu, H & Crothers, D M *Nature* **308**, 509-513 (1984)
22 Koo, H S, Wu, H-M & Crothers, D M *Nature* **320**, 501-506 (1986)
23 Tschumper, G & Carbon J *Gene* **10**, 157-166 (1980)
24 Buchman, A R & Kornberg, R, (manuscript in preparation)
25 Hieter, P, Mann, C, Snyder, M & Davis, R W *Cell* **40**, 381-392 (1985)
26 Zahn, K & Blattner, F R *EMBO J* **4**, 3605-3616 (1985)
27 Thoma, F, Bergman, L W & Simpson, R T *J molec Biol* **177**, 715-733 (1984)
28 Thoma, F & Simpson, R T *Nature* **315**, 250-252 (1985)
29 Varshavsky, A J, Sundon, O & Bohm, M *Cell* **16**, 453-466 (1979)
30 Jakobovits, E B, Brutosom, S & Aloni, Y *Nature* **285**, 263-265 (1980)
31 Saragosti, S, Moyne, G & Yaniv, M *Cell* **20**, 65-73 (1980)
32 Snyder, M & Davis, R W, (manuscript in preparation)

Flavonoid activation of nodulation genes in *Rhizobium* reversed by other compounds present in plants

J. L. Firmin, K. E. Wilson, L. Rossen & A. W. B. Johnston

John Innes Institute, Colney Lane, Norwich NR4 7UH, UK

Bacteria of the genus *Rhizobium* specifically recognize and infect legume plants, resulting in a complex morphological and biochemical differentiation into nitrogen-fixing root nodules. In *R. leguminosarum*, whose host plant is the pea, and several other *Rhizobium* species, groups of linked plasmid-borne genes are involved in nitrogen fixation (*nif* genes) and nodulation (*nod* genes)[1,2]. The *nod* genes are organized in several transcription units on the symbiotic plasmid pRL1JI of *R. leguminosarum*; these comprise *nodABCIJ*, *nodD* and *nodFE*[3–6] with *nodD* being a positively acting regulatory gene whose product activates transcription of the other two operons which contain genes essential for the nodulation process[5,7]. Although the precise biochemical functions of most of the pRL1JI *nod* genes are unknown, it has recently been discovered that their expression requires not only the product of the regulatory *nodD* gene, but also a factor (or factors) present in pea root exudate[5,7], and similar results have been reported for *Rhizobium meliloti*[8] and *Rhizobium trifolii*[9]. The finding of plant-produced products that 'signal' the *Rhizobium* offers the possibility of identifying one of the key biochemical reactions required for the infection process. Here we describe the characteristics of several plant-specified compounds that activate the transcription of the pRL1JI *nodABCIJ* and *nodFE* genes and present a novel phenomenon, namely that other, naturally occurring compounds can antagonize this induction.

For these studies the ability of different compounds to activate or to antagonize the activation of transcription of *nodABCIJ* and *nodFE* was monitored by the use of gene fusions Two *Rhizobium* test strains contained either the *nodC-lacZ* fusion plasmid pIJ1477 or the *nodF-lacZ* fusion plasmid pIJ1521, together with pIJ1518 which contains the cloned *nodD* gene[5,7] (Fig 1)

Preliminary studies showed the inducing compounds in pea root exudate to be soluble in water and methanol, dialysable and adsorbed by XAD-2 resin High-voltage paper electrophoresis showed that the active component(s) behaved as a weak acid Following high-voltage paper electrophoresis or paper chromatography (in several solvents) the region of biological activity co-migrated with material that absorbed ultraviolet light These characteristics suggested that the active component(s) could be phenolic HPLC fractionation of the exudate from germinating pea seeds showed four peaks of biological activity (Fig 2) with one (peak A) being the most prominent This was also the major peak in extracts from dried pea seeds and pea root exudate Although exudate from seeds of alfalfa, a legume not nodulated by *R leguminosarum*, could induce *nod* gene expression in this species, there were quantitative and qualitative differences in the peaks of biological activity compared with those from peas (Fig 2)

Electron impact mass spectroscopy (70 eV, 200 °C) of material in peak A from germinating pea seeds gave a complex pattern, with prominent ions at *m*/*e* 121 and 137 Although these might have been derived from simple phenolics, the presence of ions as large as *m*/*e* 256 suggested the presence of more complex phenolics One such class of compounds are flavonoids, these are made by plants and some give fragments of *m*/*e* 121 and 137

To confirm that flavonoid compounds were indeed inducers of *Rhizobium nod* genes, we took advantage of the fact that wild-type flowers of *Antirrhinum majus* contain many representatives of such molecules whereas a white flower colour mutant (termed *nivea*) which lacks chalcone synthase, contains none of these compounds[10] A methanolic extract from red wild-type flowers induced both *nodC-lacZ* and *nodF-lacZ*, whereas an extract from the white *nivea* mutant flowers did not Thus, one or more of the flavonoids in *A majus* flowers induces *nod* gene expression in *R leguminosarum* To confirm this, several commercially available flavonoids known to be present in *A majus* flowers, as well as other structurally similar flavonoids were tested as *nod* gene inducers (Table 1) Three flavones, three flavanones and one flavone glucoside induced *nod* gene transcription at low (10–100 nM) concentrations However, flavones and flavanones lacking a hydroxyl group on the B-ring had no inducer activity, vitexin (8-*C*-glucosylapigenin) and the 7-*O*-rhamnoglucosides of hesperetin and naringenin also failed to act as inducers Further, none of the isoflavones, flavanols or the flavone-3-glycoside tested had detectable inducer activity at concentrations as high as 10 μM

Two of the active compounds, eriodictyol and apigenin-7-*O*-glucoside, had identical HPLC retention times to peak A obtained from peas Evidence that peak A contained an *O*-glycoside of apigenin was that acid hydrolysis (2 M HCl, 100 °C, 2 h) of the material in peak A caused a reduction (by approximately 50%) of the inducing activity associated with this peak and generated a second, biologically active, ultraviolet-absorbing peak which had an identical retention time to authentic apigenin (Further, authentic apigenin-7-*O*-glucoside was hydrolysed using the same conditions used for the hydrolysis of the material in peak A, following separation by HPLC, it was found that all the apigenin-7-*O*-glucoside was converted to apigenin) Therefore peak A contained at least two biologically active compounds

The individual compounds tested for activation of *nodF-lacZ* and *nodC-lacZ* were also examined for their ability to antagonize or augment induction by extracts from pea seed exudate None of the compounds listed in Table 2 increased the levels of transcription of the *nodC-lacZ* or *nodF-lacZ* fusions However, it was found that, at concentrations as low as 5 μM, the isoflavones daidzein and genistein and the flavonol kaempferol strongly inhibited the activation of the *nodC-lacZ* fusion (Table 2), similar results were obtained with these compounds using the *nodF-lacZ* fusion

The second class of compounds tested as antagonists were analogues of acetophenone, which are structurally similar to the C-ring of the flavonoid *nod* gene inducers, several acetophenone analogues are known to induce transcription of the *vir*

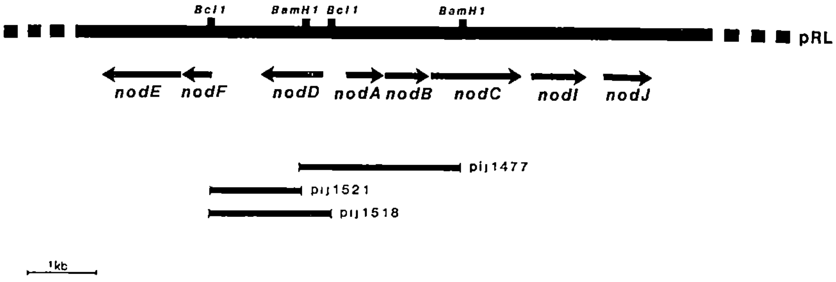


Fig 1 Location of pRL1JI *nod* genes Restriction sites used for the generation of the cloned *nodD* gene and the dimensions of recombinant plasmids used in this study The locations and directions of transcription of the eight *nod* genes are shown Restriction sites used for the generation of the cloned *nodD* gene in pIJ1518 (ref 7) and of the *nodC-lacZ* and *nodF-lacZ* fusion plasmids pIJ1477 and pIJ1521 (refs 7 and 5 respectively) are indicated

Table 1 Induction of *nodC-lacZ* expression in *R leguminosarum* by selected flavonoids

Compound	Biological activity* at flavonoid concentrations of 10 μM	1 μM	100 nM	10 nM
Flavones				
4′,5,7-trihydroxyflavone (apigenin)	5,163	4,649	541	59
3′,4′,7-trihydroxyflavone	6,814	1,069	109	64
3′,4′,5,7-tetrahydroxy-flavone (luteolin)	5,120	1,437	77	73
Flavone glucoside				
4′,5,7-trihydroxyflavone 7-*O*-glucoside (apigenin 7-*I*-glucoside)	5,964	3,642	239	55
Flavanones				
4′, 5,7-trihydroxyflavanone (naringinin)	4,552	3,154	100	58
3′,5,7-trihydroxy-4′-methoxyflavanone (hesperitin)	9,207	6,409	1,036	250
3′,4′,5,7-tetrahydroxy-flavanone (eriodictyol)	7,144	5,066	523	75

The *lac* fusion strain KH1122 pIJ1518 pIJ1477[7] was grown to a density corresponding to $A_{600} = 0\ 3$ prior to exposure to the flavonoid for 15 h at 29 °C with shaking The bacteria were then assayed for β-galactosidase activity as described previously[7] The following flavonoids had no detectable activity in the concentration range 10 μM-10 nM

Flavones 5,7-dihydroxy-, 5,7-dihydroxy, 3′,4′-dimethoxy-, 7-hydroxy-4′-methoxy-, 3,5,7-trihydroxy-, 3,4,5,7-tetrahydroxy-

Flavone glycosides 4′,5,7-trihydroxyflavone 8-*C*-glucoside, 3′,4′,5,7-tetrahydroxyflavone 3-*O*-rutinoside

Flavanone 5-hydroxyflavanone

Isoflavones 4′5,7-trihydroxy- (genistein), 4′,7′-dihydroxy

Flavanone glycosides 3′,5-dihydroxy-4′-methoxyflavanone 7-*O*-rhamnoglucoside, 4′-5-dihydroxyflavanone 7-*O*-rhamnoglucoside

* Biological activity is expressed as units of β-galactosidase activity In the absence of any added inducer the background level of expression was ~60 units

genes of *Agrobacterium tumefaciens*[11] None of these compounds induced *nod* gene expression but acetovanillone and 4-hydroxyacetophenone (at 50 μM) strongly inhibited induction of *nodC-lacZ* and *nodF-lacZ* by pea seed exudate (Table 2) At higher concentrations (1 mM) several other related phenolic compounds also strongly inhibited *nod* gene activation (Table 2) None of the compounds affected the growth rate of the cells

The results presented here show clearly that certain naturally occurring flavonoids activate *nodD*-dependent transcription of the *R. leguminosarum nodFE* and *nodABCIJ* transcriptional units and that this induction is antagonized by other flavonoids or simpler phenolics present in plants

In general, flavones and flavanones (Fig 3), which are hydroxylated in both the A- and B-rings, acted as inducers All flavonoids tested in which the C3 position was substituted with a hydroxyl group (flavonols), a sugar (flavone-3-glycoside) or phenyl group (isoflavonoids) were inactive At least one B-ring hydroxyl group appears to be necessary for inducing ability (for example 4′,5,7-trihydroxyflavone is active but 5,7-dihydroxyflavone is not) The B-ring hydroxyl can be at either (or both) the 3′ or 4′ position, thus hesperitin (3′-hydroxy-4′-methoxy), naringenin (4′-hydroxy) and eriodictyol (3′,4′-dihydroxy) are all active

The presence of a B-ring methoxy at the 4′ position appears to enhance potency, since hesperitin, the 4′-OMe-ether of eridictyol, is more active than eriodictyol itself However, no biological activity was detected when a 3′-methoxy- or 3′,4′-dimethoxy- was the only substituent on the B-ring (see Table 1) The glucosylated flavone apigenin-7-*O*-glucoside (but not apigenin-8-*C*-glucoside) acts as an inducer It is possible, though, that the 7-*O*-glycoside is deglycosylated by *Rhizobium* to form apigenin, itself an inducer of *nod* gene expression Moreover, the distinction between active and inactive flavonoids might reflect their differential uptake or metabolism rather than their (presumed) interaction with the regulatory *nodD* gene product

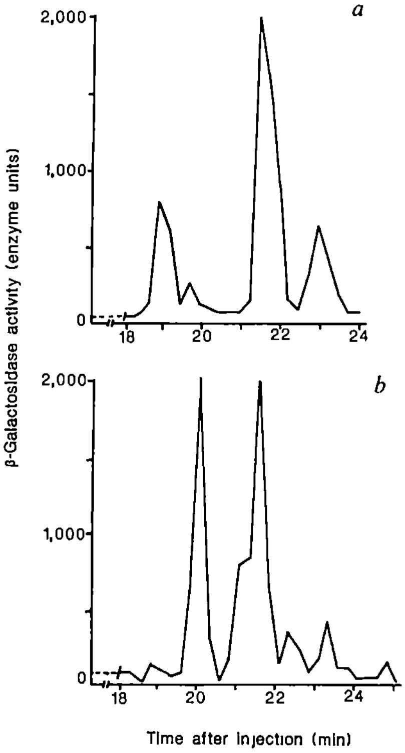


Fig 2 HPLC analysis of nod gene inducers present in the seed exudate from pea (*a*) and alfalfa (*b*)

Methods Seeds of pea or alfalfa were sterilized with aqueous sodium hypochlorite (50%) and were then soaked in sterile distilled water for 3 days with gentle shaking The liquid was removed, and after centrifugation (5,000*g*, 15 min) the supernatant was concentrated *in vacuo* (45 °C) and was applied to a preparative reverse phase (C8) HPLC column (12 cm × 0 7 cm) The column was eluted at 2 ml per min with a linear gradient of methanol-water, from 10% to 99% methanol at 3 2% per min, and fractions of 5 ml were collected After evaporation, a portion of each fraction was assayed for *nod* gene-inducing ability using *nodC-lacZ* fusion[7] Active fractions were combined, concentrated and applied to an analytical HPLC column (Brownlee Aquapore RP300, 25 cm × 0 46 cm fitted with a 3 cm × 0 46 cm guard cartridge) The column was eluted at 1 ml per min with a linear gradient of methanol-water (0-3 3 min, 10% methanol, isocratic, 3 3-24 7 min, 10-85% methanol, linear gradient, 24 7-28 1 min, 85% methanol, isocratic, 28 1-32 0 min, 85-10% methanol, linear gradient) Fractions (0 25 ml) were tested for the ability to induce *nod* gene expression as described above Similar results were obtained with the *nodF-lacZ* fusion instead of *nodC-lacZ*

Activators of *R. leguminosarum nod* gene transcription were found not only in pea, which is nodulated by this species of *Rhizobium*, but also in extracts of alfalfa, clover and cowpea plants which are not nodulated by *R leguminosarum* Thus host range specificity cannot be determined solely on the basis of *nod* gene inducers such as the active flavonoids described here

Further, the flavone luteolin was shown to be the major active component present in alfalfa seeds for the induction of *nodABC* transcription in *R meliloti*[12] This flavone also activated *R. leguminosarum nod* genes (Table 1) HPLC analysis of extracts of peas and alfalfa seed showed that they both contained several molecules with inducing activity In peas the most prominent peak on HPLC (peak A in Fig 2) contains at least two componds (this would explain the complexity of the mass spectrum) Taken together the evidence from chromatography, mass spectroscopy and acid hydrolysis strongly suggested that the material in peak

Table 2 Effect of selected flavonoids and acetophenone analogues on *nod* gene expression in *R. leguminosarum*

Compound	% Inhibition at flavonoid concentrations of 1,000 μM	500 μM	100 μM	50 μM	10 μM	5 μM
5,7-dihydroxy,3′,4′-dimethoxyflavone	ND	0	0	0	0	0
5,7-dihydroxyflavone	ND	24	20	3	0	0
4′,5-dihydroxyflavanone 7-*O*-rhamnoglucoside (naringin)	ND	0	0	0	0	0
4′,5,7-trihydroxyflavone 8-*C*-glucoside (vitexin)	ND	0	0	0	0	0
3,3′,4′,5-tetrahydroxy-7-methoxyflavone (rhamnetin)	ND	83	68	30	0	0
3,4′,5,7-tetrahydroxy-flavone (kaempferol)	ND	82	82	65	69	60
4′,7-dihydroxyisoflavone (daidzein)	ND	85	88	89	85	75
4′,5,7-trihydroxyiso-flavone (genistein)	ND	91	82	58	78	81
4-hydroxy-3,5-dimethoxybenzoic acid (syringic acid)	83	39	31	29	23	ND
4-hydroxy-3,5-dimethoxy-benzal-dehyde (syringal-dehyde)	88	37	22	23	23	ND
4-hydroxy-3,5-dimethoxycinnamic acid (sinapinic acid)	89	60	28	15	22	ND
4-hydroxy-3-methoxyacetophenone (acetovanillone)	95	95	91	85	20	ND
4-hydroxy-3,5-dimethoxyaceto-phenone (acetosyringone)	96	94	95	46	5	ND
3,5 dimethoxyaceto-phenone	92	69	14	21	19	ND
4-hydroxyaceto- phenone	94	95	94	91	41	ND

The *lac* fusion strain KH1122 pIJ1518 pIJ1477[7] was pregrown to an absorbance (at 600 nm) of 0.15 prior to exposure to the compound under test, in the presence cf a fixed amount of natural inducer (obtained from germinating pea seeds) equivalent to 2,000 units of β-galactosidase activity. The bacteria were shaken for 15 h at 29 °C and were then assayed for β-galactosidase activity as described previously[7]. ND, not determined.

A contained apigenin-7-*O*-glucoside plus another flavonoid which had the same chromatographic behaviour as eriodictyol. Further, Redmond *et al.*[13] found that in clover the major activator of *R. trifolii nod* gene transcription was 7,4′-dihydroxy-flavone.

The finding that other naturally occurring compounds made by plants inhibit *nod* gene induction by flavonoids is novel. Those flavonoids which were most potent in antagonizing induction had hydroxylation patterns in the A- and B-rings similar to those of the inducing flavonoids described here but different from the inducing molecules in having a substituted C3 position, as in the isoflavonoids genistein and daidzein and the flavonol kaempferol. Flavones lacking a B-ring hydroxyl (for example, 5,7-dihydroxyflavone) were much less inhibitory. Given the structural similarities between antagonists and the inducing molecules it is possible that they act competitively either at the level of transport or within the bacterial cell. The substitution pattern of the acetophenone analogues tested also had a marked effect on their potency as inhibitors of *nod* gene induction, the most potent of this group were those which had a substitution pattern [(3-methoxy, 4-hydroxy-) or 4-hydroxy-] similar to that found in the B-ring of the flavonoids that induced *nod* gene transcription.

There is an apparent paradox in the fact that isoflavonoids were among the most potent antagonists of *nod* gene transcription, *a priori*, it might have been thought that isoflavonoids,

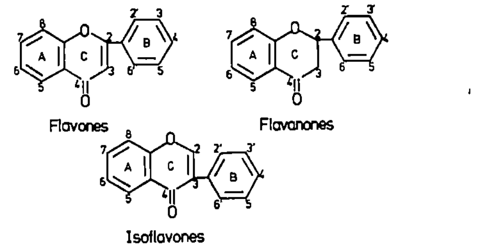


Fig. 3 Skeleton and numbering nomenclature for the flavonoids mentioned in this study.

which are found predominantly in the family Leguminosae[14], would be good candidates for compounds that activate the expression of *Rhizobium nod* genes.

The concentrations required for antagonism of induction by, for example, the isoflavonoid daidzein were one to two orders of magnitude greater than the concentration required for *nod* gene induction by the most potent of the activating molecules. However, isoflavonoids are present in germinating legumes at concentrations as high as 20 μM even in healthy plants[15]. It seems likely therefore that, during the normal nodulation process, the regulation of *Rhizobium nod* genes is subject both to activation and to the antagonism of activation by different plant-specified molecules. Indeed, compounds were found in pea root exudate which antagonize the induction of the *nodC-lacZ* fusion by the inducing compounds in peak A. This was demonstrated by fractionating root exudate by the HPLC and adding individual fractions of exudate to a strain carrying a *nodC-lacZ* fusion which had been grown in the presence of the inducer (C. A. Shearman, personal communication). The host plant legume can regulate the number of nodules it contains on its root system. This has been clearly demonstrated by the fact that mutants of soybean[16] and of pea[17] have been isolated which contain many times more nodules than wild-type. Further, when a recombinant plasmid containing the cloned pRL1JI *nodABC* genes, transcribed constitutively from a vector promoter, was introduced into a wild-type strain of *R. leguminosarum* nodulation of peas was severely inhibited[18] indicating that continuous high-level transcription of these genes is deleterious for the infection process. It will be of interest to see if the control of nodule number depends upon the extent to which the *nod* genes are expressed, this being determined by the relative concentrations of the inducers and anti-inducers which are present in the rhizosphere and in the plant root tissue.

We thank Keith Parsley for mass spectrometry, Enrico Coen for providing *Antirrhinum* plants and David A. Hopwood for useful discussions about the manuscript.

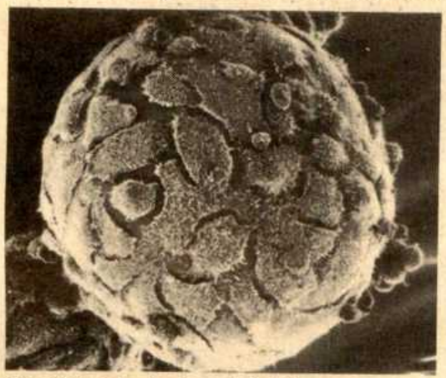
Received 23 May; accepted 14 July 1986.

1. Long, S. R. in *Plant Microbe Interactions, Vol. 1, Molecular and Genetic Aspects* (eds Kosage, T. & Nester, E. W.) (Macmillan, New York, 1984).
2. Rossen, L. *et al.* in *Oxford Survey of Plant Molecular and Cell Biology* (ed. Miflin, B. J.) (in the press).
3. Downie, J. A., Knight, C. D., Johnston, A. W. B. & Rossen, L. *Molec. gen. Genet.* **198**, 255-262 (1985).
4. Rossen, L., Johnston, A. W. B. & Downie, J. A. *Nucleic Acids Res.* **12**, 9497-9508 (1984).
5. Shearman, C. A., Rossen, L., Johnston, A. W. B. & Downie, J. A. *EMBO J.* **5**, 647-652 (1986).
6. Evans, I. J. & Downie, J. A. *Gene* (in the press).
7. Rossen, L., Shearman, C. A., Johnston, A. W. B. & Downie, J. A. *EMBO J.* **4**, 3369-3373 (1985).
8. Mulligan, J. T. & Long, S. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6609-6613 (1985).
9. Innes, R. W. *et al.* *Molec. gen. Genet.* **201**, 426-432 (1985).
10. Spinbille, R. & Forkmann, G. *Phytochemistry* **21**, 2231-2234 (1982).
11. Stachell, E. S., Messens, E., van Montague, M. & Zambryski, P. *Nature* **318**, 624-629 (1985).
12. Peters, G. M. & Long, S. R. *J. Bact.* (in the press).
13. Redmond, J. W. *et al.* *Nature* (in the press).
14. Harborne, J. B. in *Chemotaxonomy of the Leguminosae* (eds Harborne, J. B., Boulter, D. & Turner, B. L.) 31-71 (Academic, London, 1971).
15. Partridge, J. E. & Keen, N. T. *Phytopathology* **66**, 426-429 (1976).
16. Carrol, B. J., McNeil, D. L. & Gresshoff, P. M. *Pl. Physiol.* **78**, 34-40 (1985).
17. Jacobsen, E. & Feenstra, W. J. *Pl. Sci. Lett.* **33**, 337-344 (1984).
18. Knight, C. D., Rossen, L., Robertson, J. G., Wells, B. & Downie, J. A. *J. Bact.* **166**, 552-558 (1986).

Notes for the laboratory

From cell culture products to an expert program, this week highlights innovative and useful ideas.

Ventrex Laboratories, Inc. has recently introduced **neutrally buoyant glass microcarrier beads** (Reader Service No. 100). Typical Ventreglas beads have a specific gravity of only 1.03 making the vigorous stirring speeds that retard cell growth, but are required with heavier beads, unnecessary. Ventrex's beads are polystyrene coated with glass, and are therefore capable of withstanding cleaning by trypsinization, chromic acid, or alcohol. Ventreglas beads are also autoclavable. The beads come in a variety of diameters ranging from 90–210 µm, and a 25 g sample is available for $100 (US).



The Ventreglas bead: unsinkable

Statistical Programs' new expert computer system, EXPERtIMENTAL DESIGN, eases the selection of **statistical designs for data analysis** (Reader Service No. 101). Through a question-and-answer format, the program ranks the suitability of 13 designs, ranging from factorial to Box-Behnken analysis, to the user's experiment. Uncertainty in the user's answers is dealt with by assigning desirabilities and manipulating them according to fuzzy set theory. Accompanied by a user's manual including a tutorial, the programmed floppy diskette sells for $295 (US), and is written for 256K IBM PC/XT/AT.

Mycoplasma contamination need not plague your cell cultures and influence your experimental results, now that the American Type Culture Collection is offering their new **mycoplasma detection service** (Reader Service No. 102). ATCC's procedure costs $50 (US) each for 1–10 samples, and utilizes both a direct and an indirect assay to determine if suspect cultures are affected. The direct assay involves a total incubation time of 28 days, according to ATCC, while the indirect assay (using a fluorochrome staining procedure) can be accomplished within a week.

The membrane-based **culture plate inserts** for epithelial cell culture available from Millipore allow *in vitro* cultures to closely mimic cells grown *in vivo* by providing cells with both serosal and mucosal contact with media (Reader Service No. 103). Millicell-HA inserts come presterilized in blister packages for convenience, and are produced in a 12 mm configuration to fit 24-well plates, as well as a 30 mm configuration appropriate for 6-well plates. Available since June 1986 in the United States, Millicell-HA inserts are now available in the United Kingdom, and are priced at £50–70 (UK) (depending on size) for a box of 50.

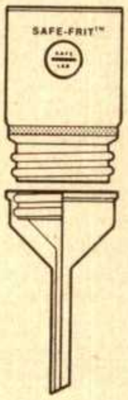
Taconic, Inc. eliminates worries over **clinically silent infections in laboratory animals** with its free bimonthly *Lab Animal Health Report* (Reader Service No. 104). Each edition will feature a different strain, and will outline the battery of tests to which Taconic mice and rats are subjected, along with ratios of incidence of infection. Available to any researcher using mice and rats, Taconic's reports will allow experimenters to exclude subclinical infection from other influential variables in the interpretation of results. ☐

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENTS

SAFE-FRIT™ FUNNELS

Safe-Lab's NEW fritted glass funnels feature a removable, unbreakable polypropylene stem just below the fritted glass disc. This unique design eliminates the danger of broken glass funnel tips and permits easy cleaning from both sides. Removable stem allows you to set the funnel top on counter or directly inside an oven or desiccator for drying, without the danger of tipping or spilling. Polypropylene has excellent chemical resistance and threads will not leak, even at high-vacuum.



Call or write for a FREE catalog:
Safe-Lab, Inc., P.O. Box 1290,
Santee, CA 92071
(619)448-9100

Reader Service No.28

MICRO-TOOL MAKING!
then you need the
SINGER MICROFORGE

Write to: **Singer Instrument Co Ltd, Treborough Lodge, Roadwater, Watchet, Somerset TA23 0QL, England. Tel: (0984) 40226. Telex. 337492 Comcab G.**

Reader Service No.50

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

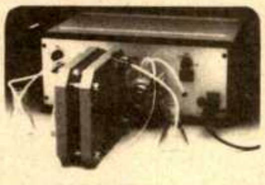
Low Cost

Eliminates evaporation loss and contamination. Easy to use – flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. **billups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.**

Reader Service No.7

LIPOSOMES
REPRODUCIBLE – HOMOGENEOUS – EXTREMELY RAPID

NEW



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM
POB 126, D-8000 Munich 65, FRG

Reader Service No.35

The changing graduate labour market

Richard Pearson

Competition between employers for the best UK graduates is becoming intense.

As the new academic year gets under way, final-year undergraduates have to face up to the twin pressures of final exams and job hunting. While 1987 promises to be another bouyant year for graduate recruitment, the traditional pattern of the annual 'milkround' is coming under increasing strain as employers seek the 'best' graduates at a time when the universities are producing fewer graduates.

Thus, alongside generally booming job prospects, the graduate labour market is showing clear evidence of fragmentation, reflecting a more general trend in the labour market as the market becomes 'freer' and employers compete intensively for the 'best' and for specialist skills, offering increasing salary differentials and special premiums for quality staff, while those with less marketable skills suffer at the bottom end of the labour market chasing low paid and often temporary jobs, or experiencing bouts of unemployment. As a result, while 1985 saw the largest ever number of university graduates entering employment, unemployment six months after graduation remained at significant levels, 8.6 per cent, although this was somewhat lower than in previous years (Table 1) and compares favourably with unemployment in the population at large.

This year well over 100,000 students graduated with first degrees, with the universities accounting for two out of three graduates, and the polytechnics and colleges one in three. The majority graduated in arts and social science subjects, with under 15 per cent being engineers and technologists, while a further 28 per cent were scientists (Fig. 1). Not all of those graduating, however, seek to, or actually enter the labour market, about one in four going on to further academic work or formal training (although only 3,200 went into teacher training in 1985, a fall of over 40 per cent since 1981) while about one in six were unemployed or otherwise not available for work six months after graduation. The 33,500 university graduates going into employment were split in roughly equal proportions between public service/education, industry, and finance/commerce, with the latter having shown the most growth during the past five years.

Table 1 Graduate unemployment

	Unemployment six months after graduation (%)	Entering permanent employment
1979–80	9.0	29,882
1980–81	11.3	28,523
1981–82	13.5	29,602
1982–83	11.6	32,506
1983–84	9.6	33,942
1984–85	8.6	34,717

Source: University Grants Committee

Interestingly, while the numbers entering jobs in industry had grown steadily between 1980 and 1984, they fell slightly in 1985. For this year the forecasts have been of a five per cent increase in overall vacancies with particular growth being shown in the fields of transport, government, public utilities, financial and other services. The decline in engineering and other manufacturing vacancies is expected to continue.

Traditionally the vast majority of graduate vacancies have been advertised and filled through the annual 'milkround', when employers visit the campuses in the spring term, announcing their vacancies and carrying out first interviews. Nowadays, however, an increasing number of employers are starting their visits before Christmas and making job offers much sooner in order to be able to select the best staff. The numbers doing this have risen from 12 to 17 per cent of recruiters in the last three years (*The Cost Effectiveness of the Graduate Milk Round*, IMS, 1986). At the same time more and more vacancies are being advertized through the summer fairs, with over one in three recruiters using them in 1986, some to compensate for shortfalls from the milkround but also because some new vacancies appear in the summer months. More students also appear to be delaying job-searching to concentrate on their exams, secure in the knowledge that good jobs will be available in the summer months. Thus while the milkround remains the dominant recruitment method, the graduate labour market is becoming less concentrated in time.

Another sign of the fragmentation of this part of the labour market is the increasing range of starting salaries being offered. In the early 1970s virtually all graduates received similar starting salaries with only small distinctions being made between different types of employers and little distinction being drawn between graduates from different disciplines. By the late 1970s, premiums were being paid to engineers in shortage disciplines, and with the demise of incomes policies, more emphasis on market forces and increased competition, the diversity of starting salaries has increased markedly. For example, in 1986 the median graduate starting salary was about £8,000 per year, but with an upper decile of £9,000 and lower decile of £6,750. The annual rate of increase at the top of the range of 9 per cent was double that at the bottom, further widening the gap.

Last week's 'Big Bang' in the City of London had also had a highly publicised impact on graduate salaries, with 1986 seeing many examples of new city recruiters offering starting salaries of between £10,000 and £13,000 per year — in several cases even more. Looking to 1987 it seems clear that many more financial institutions will be offering starting salaries in between £12,000 and £14,000 a year, which will be well ahead of the likely average of about £9,000 for most jobs. With golden hellos also being paid (one merchant bank, for example, giving 'bursaries' of £2,000 to let graduate recruits 'travel in the summer vacation'), decisions about jobs and careers can become even more complex.

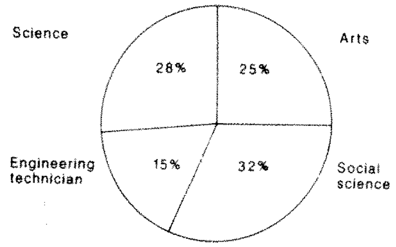


Fig. 1 Subject distribution, all institutions. (Source: University Grants Committee.)

The key for job hunters is to look beyond the short-term rewards of salary and also think about job interest, training and experience, and the longer-term career development prospects, which may be gained through several firms. While these latter factors can often act as compensation for a lower starting salary, the disturbing feature of many jobs on offer is that the lower salaries are often combined with the poorer training and career prospects, further emphasizing the fragmentation of the labour market into a smaller elite of highly rewarded and well trained staff, pulling away from the broader range of jobs available to the majority, and unemployment or underemployment facing a significant minority. ☐

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton BN1 9RF, UK.

nature

What price independence?

The British government has scraped together a little money to save its universities from catastrophe, winning in return a series of promises that academics may not be able or willing to keep.

BRITISH universities cannot afford to decline last week's offer of extra funds (see p. 97), but they will pay a high price for only a little extra. Not that the strings attached to the offer by Mr Kenneth Baker, the new Secretary of State for Education and Science, are all that onerous. He demands selectivity in research, the elimination of small departments, better management, the abolition of academic tenure, more technology transfer to industry, more explicit concern about academic standards and more help with the supply of mathematics and physics teachers in the schools. What could be more reasonable? Unfortunately, there is a catch.

What Mr Baker asks is not new. Some of his requests are even matters on which the university system itself has taken the initiative. Selectivity, for example, has been a slogan for at least twenty years. More recently, it has been the practical response of the University Grants Committee (UGC) to the shrinkage of its budget; its intention to concentrate part of its funds on institutions with a reputation for research is two years old, but began painfully to bite with this academic year. Similarly, for better or worse, British universities are now thoroughly steeped in the doctrines on management put out last year by what is called (after the chairman of a committee) the Jarratt report, commissioned jointly by the Committee of Vice-Chancellors and Principals and UGC. So why fear the gift of extra funds when the strings attached require largely the pursuit of present goals?

The obvious danger is that the prudent intentions of titularly autonomous institutions are turned into external obligations that can (and will) be monitored. Selectivity? A single civil servant in his spare time will be able to inspect the disparities in the treatment by UGC of individual universities and complain if they are too small. Better management? UGC has undertaken to carry out the first invigilation itself. Academic standards? The vice-chancellors, after what has evidently been a busy summer, have put out a mixed dossier of codes of practice (on matters such as the training of academic staff and the role of external examiners) and of half-considered means of assessing academic performance (citation counts, for example), giving further hostages to the external monitors. The vice-chancellors have even promised to produce a report on present practices of awarding academic tenure; without promising to abolish tenure, the vice-chancellors' letter promises to "aim" at ensuring that people "are not initially appointed with tenure in its strongest form where this can be avoided".

Accountability

But should not universities supported by public funds be publicly accountable? Of course. The danger in the new relationship between the British government and its universities is that the social and economic value of the academic system will be judged by criteria devised for managerial purposes only. It should be especially galling for academics, and alarming as well, that Baker was able to attach to last week's written statement copies of letters from the chairmen of UGC and of the vice-chancellors' committee embodying written promises that they will conduct themselves appropriately in the future.

How could that have come about? In the idiom of British public life, the circumstances are easily reconstructed. Baker differs from Sir Keith Joseph, his predecessor, in preferring to make friends rather than enemies. He is also adroit, so much so that the failure of his first project for the present government, the reinvigoration of the British economy by the encouragement of cable television, has left his reputation unscarred. At some early meeting with the chairmen of UGC and the vice-chancellors' committee, he would have explained that he personally is sympathetic to their difficulties (and appalled at the prospect that one or more universities might have to close), would have added that the government in which he shares collective responsibility is unfortunately convinced that academics are an expensive bunch of unproductive layabouts, would have listened sympathetically to the universities' account of how strenuously they have tried to put their own houses in order (and would generously have advised them to spend more on making their efforts better known), would have offered to fight hard for a little extra cash in the run-up to the annual settlement (last week) of next year's public spending and, pleading that he himself would be occupied trying to bring the semblance of indusrial peace to British schools, would have asked that academics should help him and their own cause by providing the ammunition he would need.

Contrition

The sting has worked. There will be a little extra money. Baker's civil servants would have provided guidance (in the British idiom) about the kind of ammunition they would need (but their first proposal for a tripartite statement appears to have foundered, improbably at a gathering of academics in Penang). But Baker has been able to deliver to his colleagues a prize that eluded his predecessor — a show of contrition for imagined misdemeanours by the universities, a promise to do better in future and a set of measuring-rods with which they may be beaten if the occasion arise.

Nothing that happened last week touches the true problem of the British university system, which is structural and neither financial nor managerial. The universities boast that they are autonomous, but that is less than half the truth. In reality, they are public corporations almost totally dependent on public funds (on UGC and on local education authorities for their teaching functions, on the research councils for research). In the halcyon 1960s, when the only problem was that of investing quickly enough in higher education, UGC's performance easily upheld the British conceit that it is a uniquely judicious way of channelling public funds towards private ends.

The novelty, now, is that British governments, having become by accident providers of higher education, have also for some time become disenchanted with the product (or at least persuaded that higher education can be had more cheaply in other kinds of institutions than universities). UGC, no longer able to act the benefactor, has taken to telling particular institutions that they cannot teach this, or that ... or else. One of UGC's latest projects, an enquiry into the viability of "small" departments of physics and chemistry, will almost certainly discover that they are financially uneconomic and will recommend that they should be abolished or amalgamated. UGC will regret, like the institutions concerned, that universities thus deprived will hardly meet the classical concept of the university.

Especially under its present enlightened management, there is no reason to believe that UGC wants to become what it is well on the way to being — every British academic's dean's office. Nor can Baker seek the kind of centralization now being forced on the system by financial stringency. He should know that there is a respectable way out, one that would get the government out of the business of setting the specific goals by which universities conduct themselves. He would be nudged in that direction if the committee now looking at UGC's own function, due to report in the next twelve months, recommends regretfully that UGC should be abolished.

But how could British universities survive without UGC? By being challenged with the offer of the autonomy of which they boast. It is now five years since they were told (by UGC) that they would be penalized financially if they exceeded certain student numbers; the threat of penalty has been lifted, but target student numbers remain the basis on which UGC allocations are calculated. One result is that successful universities cannot grow as quickly as they might, while less successful institutions are kept in being, at least in the short run, by continuing subventions. The prospect, as UGC enforces rationalization for sound economic reasons, is that many of these institutions will become lop-sided places, covering only part of the ground of scholarship and thus even less suitable for, and probably even less attractive to, potential students. The truth is that British universities are, on the average, too small for the smaller among them to be complete. Their size probably makes them uneconomic as well. The general interest would be served by amalgamation, for which, under present arrangements, there is no incentive.

But how could autonomy be afforded? Curiously, the financial problem is not nearly as insurmountable as it may seem. A third of the UGC budget is reckoned to cover research support, leaving £1,000 million a year in round numbers for the support of other academic functions in universities. Public funds are channelled to institutions of higher education in the form of tuition fees fixed far below average or even marginal cost. Although local authorities administer the system, central government meets the cost. So why not allow tuition fees to rise again (they were arbitrarily cut four years ago when Baker's predecessor feared they would partly protect some universities from the penalties for quota-breaking), meeting the cost from the central budget of UGC? If managed over a predetermined period of time, and coupled with the understanding that students would choose between universities, the result would be both autonomy and a measure of competition between institutions at present persuaded that they must all sink together.

But, Baker's civil servants will ask, what assurance will there then be that the universities are publicly accountable? How will it be possible to tell that academics are indeed carrying out a modicum of research, or that they are taking full account of citation counts in promoting their own members to more senior posts? These questions are of course irrelevant. If the universities were financially autonomous, dependent on public funds only indirectly through the preferences of students for courses of study and for places at which to follow them, the universities would have an incentive which they at present lack to manage their affairs efficiently. They would also have an incentive towards diversity, of which Britain has too little. The endless quarrel with the government about tenure as universities came to put their institutional survival before the interests of individuals would end. Some might even discover that they need not stick slavishly to UGC "norms" for the number of students per member of teaching staff any more than they are compelled to follow UGC standards for the equipment of buildings. To be sure, in such a regime, only the better universities would prosper. Is there anything wrong with that? □

Research in trouble

Mr Baker's little extra has also created a dilemma in which leaving CERN may be the only choice.

THE research half of the British academic enterprise will be hardly better pleased than the university part with last week's events. Outwardly, the research councils will have an extra £39 million to spend next year, over and above inflation, or more than 6 per cent of total public support for basic research. One immediate snag is that £14 million of the extra offered last week was also offered just under a year ago by Sir Keith Joseph, and has long since been counted on by the Advisory Board for the Research Councils (ABRC) on whose advice the sum available is eventually divided among the five research councils. Another and more serious difficulty is that the British basic research enterprise is yet again in the throes of a crisis thrust upon it by circumstances over which it has no control. Sterling has once again depreciated, with the result that the Science and Engineering Research Council (SERC), which must find from its budget international subscriptions to organizations such as CERN, the European high-energy physics laboratory, is short of roughly £20 million, nearly a quarter of what it spends each year on research grants to academics. ABRC will have a stark choice at its two meetings later this month: either give the whole of Mr Baker's extra money to the research council in most urgent need of it, letting the others smoulder in penury for another year, or invite SERC to face a decision that has been staring it in the face for the past three years, whether to pull out of CERN.

It is a curious state of affairs. Except for 1980, when sterling appreciated against European currencies (and SERC's predecessor spent its windfall gain on extra equipment grants for universities), the crisis is chronic. Nobody is astonished, for sterling is a steadily depreciating currency. To saddle the research councils with this perpetual difficulty is of course unfair; while the international organizations to which SERC subscribes have scientific objectives, the government also has a political interest in their continuation. With its new policy on space research, the British government would not readily agree to a decision of SERC to save its annual subscription to the European Space Agency by pulling out of that. British membership of CERN is a political matter for different reasons; the British government has an interest in seeming no less faithful a European state than its competitors. But, with the exception of a single year when the Treasury was persuaded to do the decent thing, there has been no occasion in the past decade when the research councils have been adequately bailed out. Members of the council's peer review committees meeting in the past few weeks to consider new grant application have been startled to be told that SERC is £20 million short, for which reason it may be necessary to postpone the dates at which some successful applications become effective. Of necessity, the grants postponed will be those that can be put off practically; they may well include the more promising projects, which is no way to run a research enterprise already low on morale.

This is why withdrawal from CERN is yet again a live issue. ABRC's own inquiry under Sir John Kendrew concluded that it would be better, first, to negotiate a reduced budget (or British contribution), which is now the purpose of a study by CERN members due to be completed next year. It would be discourteous to pull out before that review is completed, but SERC may have no choice. The way things have turned out in Britain, with the research community persuaded that the future is only bleak, the decision may not matter much. But if there is a chance of rescuing something from the British enterprise, it is more important to give a few young people with bright ideas a chance to turn them into reality than to continue membership of even such a splendid organization as CERN. But if SERC goes down that road, it should promise itself and its constituents that it will also look hard at the need for all of its own domestic facilities. □

British penury relaxed

Increases for universities just avert disaster

BRITISH universities and the research councils are to benefit from a modest cash injection in the two years from next April. This is their share of a £4,000 million package of spending increases unveiled last week by the government, and widely taken to be part of the preparation for the next general election, which must be held before the spring of 1988.

Although education as a whole receives the lion's share of the whole package, which includes £460 million for the improvement of the pay of teachers in schools, the universities are unimpressed with their share. This amounts to an extra £61 million in 1987–88 and an extra £71 million the following year. The research councils are to get an extra £39 million a year, beginning with the next financial year next April.

The Committee of Vice-Chancellors and Principals (CVCP) welcomes the extra funds as anybody would welcome "water after years of drought", saying that the new budget will do little more than avert disaster. CVCP was especially indignant at the meagreness of the increase for the research councils, saying that the sum is not only disappointing but that it does not compare with the investment in research by Britain's industrial competitors.

Announcing the increases last week, Mr Kenneth Baker, Secretary of State for Education and Science, more than hinted that there are strings attached. In particular, the university increases are dependent on continuing progress in the rationalization of research and the channelling of sufficient funds into applied research.

Baker, who succeeded Sir Keith Joseph as secretary of state only in June, said last week that he "would expect" the University Grants Committee (UGC) to provide for some early retirement and the recruitment of new young academic staff and special initiatives in teacher training to help deal with the shortage of teachers of mathematics, physics and technology in British schools.

The execution of this policy is in the hands of UGC, which began the process of selectivity last May, when universities were told individually that public support for this academic year would for the first time be linked with research performance. The equity of that unequal distribution of funds was widely disputed (see *Nature* **322**, 299; 1986); by the end of next year UGC and the research councils are to produce a joint statement of the criteria that will be followed in the further pursuit of selectivity in research.

Kenneth Baker's cash has strings attached.

Meanwhile, UGC says that whatever the funds available through the research councils, "universities separately need to draw up and carry out plans for the selective distribution of the resources they devote to research". This will lead to "better research through the concentration of effort in strong departments".

British universities are also under continuing pressure to improve their internal management along lines recommended by the Jarratt committee in 1985; responses by individual universities are now in the hands of UGC. The government is also pressing for common criteria of academic quality and the abolition of academic tenure.

UGC is responding to the further pressure to rationalize the pattern of university activity in Britain by a series of reviews of academic areas. Reviews have already been completed of teaching and research in pharmacy, agriculture, oceanography, Italian and Scandinavian studies.

There will be further pressure on the university system later this month, with the planned publication of a Cabinet Office report on publicly supported research. The government has not hidden its discontent with the efficiency with which research is transferred from universities and government research establishments to the commercial sector. **Bill Johnstone**

University Grants Committee grants (£ million, year starts 1 April)

	1986–87	1987–88	1988–89
Previous plans	1,318 (108)*	1,356 (106)	1,392 (110)
November 1986 plans	—	1,413 (110)	1,463 (110)
Increase over previous plans	—	57 (4)	71 (0)
		Total: 61	Total: 71

* Values are for the recurrent grant, with equipment grant given in parentheses.

Meningitis

Puzzling cluster in rural Britain

THE incidence of meningitis in the small English parish of Stonehouse in Gloucestershire, which is 14 times the British average, has now prompted the British Department of Health to mount a special investigation. Forty doctors, nurses and scientists will shortly complete the first phase of a programme to try to determine the reason. It is hoped that the 7,000 residents of Stonehouse will agree to be tested.

The subject of the study is principally a bacterial strain of meningococcal meningitis (B15R) which is resistant to Group B Type 15 sulphonamide. In the past five years, 53 cases of meningitis, mostly B15R, have been diagnosed in the Stroud area of Gloucestershire (population 103,600), which includes Stonehouse. Since 1984, Stonehouse has recorded 10 out of 12 diagnosed cases of the disease from among the 2,000 people living on one housing estate. The national average for the incidence of meningitis is 1 per 100,000 of the population. The incidence in the Gloucester district is five times that figure, and that of Stroud is double again with 9 cases per 100,000.

The research project is being coordinated by Gloucester Public Health Laboratory and the Gloucester Health Authority with support from the Department of Health and Social Security (£57,000) and the Meningitis Trust (£20,000), a charity recently established in the Gloucester area to fight the disease.

The volunteers in the programme will submit themselves to swabs and blood and saliva tests. The results of the first two categories of data are expected to identify the carriers of the disease and those with immunities; it is hoped that the other, conducted by a team from the University of Edinburgh, will indicate the type of person susceptible to contracting it. The Edinburgh team hopes to substantiate its theory that those who do not secrete blood group substances into their saliva are more likely to develop meningitis.

Although the disease can be successfully treated with antibiotics if diagnosed early, some young victims have died within two days. Fever, rash, vomiting and eventual coma are the classic symptoms. But although the under-fives are more susceptible to the ordinary form of the disease, the B15R strain is claiming many of its victims from among young adults, most of them teenagers.

The numbers of diagnosed cases of meningitis peaked in 1974 with 1,296 cases and dropped to an all-time low ten years later with 404. But last year the figure was 549 and 900 cases are expected this year.

Bill Johnstone

US mid-term elections

Power jockeys in new race

Washington

Last week's elections put the Democratic Party front and centre on the Capitol Hill stage. With control of both the Senate and the House of Representatives, the Democrats will carve up the spoils of their victory in the next few weeks, chiefly the chairmanships of the committees that determine the legislative agenda for the next session of Congress. But a new majority party will hardly be a panacea for budget troubles, and the Republicans will now have the luxury of watching the Democrats make the difficult choices.

The Democrats gained control of the Senate, and now enjoy a 55–45 majority. Several extremely close races delayed final calculations, but the Democrats will add at least 5 seats to their 253–182 majority in the House of Representatives. Ironically, several of the Republicans being replaced as Senate committee chairmen are reckoned more liberal than their Democrat successors. An example is John Stennis, a conservative southern Democrat, who takes control of the powerful Appropriations Committee from Mark Hatfield, a liberal Republican from Oregon.

The chairmanships of the key appropriations subcommittees controlling the purse strings for the National Science Foundation (NSF) the National Institutes of Health (NIH) and other government science agencies are still uncertain, but present chairmen have treated science agencies generously. Lowell Weicker (Republican, Connecticut) has been a strong ally of NIH as chairman of the health subcommittee, and Utah Republican Jake Garn has been favourably disposed toward NASA (the National Aeronautics and Space Administration) as chairman of NASA's subcommittee. Garn was a passenger aboard the space shuttle in 1985.

The Strategic Defense Initiative (SDI), already under attack in Congress, will probably face greater resistance from the Democrats taking over military and energy committees. Both Sam Nunn of Georgia and Louisiana Democrat J. Bennett Johnston have criticized SDI.

A dramatic philosophy shift will occur on the Labor and Human Resources Committee when Orrin Hatch of Utah steps down in favour of Edward Kennedy of Massachusetts. Hatch, a conservative, has been an outspoken opponent of abortion, favouring a constitutional amendment making it illegal. Kennedy is a liberal and a strong proponent of government-supported health-care programmes.

In the House of Representatives, membership of science committees is unlikely to change substantially. The most significant difference will be the departure of Don Fuqua from the House Science and Technology Committee. Fuqua, who retired this year and will take a job as a lobbyist for the space industry, will be replaced by Robert Roe of New Jersey.

Although science agencies fared relatively well in the most recent budget cycle, next year will probably be a different story. The Gramm–Rudman deficit reduction act will force still larger cuts to reduce the budget deficit.

William Wells of George Washington University, who follows congressional science policy for the American Association for the Advancement of Science, says a harbinger of next year's problems was seen in the House of Representatives appropriation for the NSF research budget. Although NSF has received favourable treatment from the Democrat-controlled House, this year the appropriation for NSF-sponsored research was $146 million less than the administration request.

With only a fixed amount of money to dispense, the House subcommittee chose to protect certain housing and environmental programmes at the expense of NSF. In the final budget compromise with the Senate some of that money was restored, and overall NSF received about an 8 per cent increase in funds. According to Wells, in the next Congress just holding steady will be a "a remarkable achievement". **Joseph Palca**

British SDI cash

The value of SDI contracts in Britain has not so far approached the reported $1,500 million target sought by British politicians at the outset, but their total value now exceeds $14 million. Contracts and subcontracts include, according to the Federation of American Scientists and the Council on Economic Priorities, Barr and Stroud $250,000; Cap Scientific $250,000; British Aerospace $1.25 million; Ferranti $508,000; General Electric (UK), $1.75 million; Hunting Engineering $1 million; Logica $700,000; PA Defence Services $750,000; Plessey $1.45 million; Racal $250,000; Scicon $250,000; Short Brothers $250,000; Thorn EMI $500,000; UK Atomic Energy Authority $4.3 million; UK Ministry of Defence $1.4 million.

The Science and Engineering Research Council's Rutherford Appleton Laboratory is also now negotiating a $15 million research contract for the Super-Sprite Laser Facility. In the British Isles those doing SDI work include University College, Cork, Heriot-Watt University and the University of London. Tim Beardsley

---

Strategic defence

SDI should change course

Washington

Research for the Strategic Defense Initiative (SDI) continues to be an obstacle to superpower arms negotiations, but a compromise on defensive research that meets the legitimate interests of both sides should be possible, according to the Aspen Strategy Group, an independent think-tank.

The group's study, *The Strategic Defense Initiative and American Security*, published on 22 October, is critical of the current SDI programme, saying its large-scale mockups are inadvisable for technical reasons and erode confidence in the Anti-Ballistic Missile (ABM) Treaty. But the Aspen group says "progress towards a new [strategic arms] agreement should be possible" if the United States modifies its position that all development and testing of space-based defensive systems is allowed, in return for the Soviet Union dropping its demand that only laboratory research is possible.

The co-chairmen of the Aspen study were William J. Perry, a former undersecretary of defence under President Carter, and Lt-Gen. Brent Scowcroft, national security adviser to President Ford. The group includes senior politicians, academics and representatives of industry.

The group sees valuable opportunities to trade arms cuts for a commitment not to give notice of withdrawing from the ABM treaty for 10 years — this would mean forfeiting "very little in technical terms". The study finds "virtually no prospect" of a significant and effective population shield against a responsive enemy this century. And it says the SDI programme's large field experiments "appear to be driven by political considerations" and risk "freezing in" technologies prematurely in what should be a long-term programme.

The Aspen group also says the United States has conducted SDI experiments "whose consistency with [the ABM treaty] is measured by criteria that we would probably reject if the Soviets used them to justify their programmes". The US administration has proposed a new interpretation of the treaty that would allow testing of SDI components as long as they were not actual weapons.

The group recommends that SDI research should be refocused on more long-term "high-leverage" goals, including the problem of mid-course discrimination of warheads from decoys and that defensive research should have a near-term goal of preventing or responding to a Soviet break-out of the ABM treaty and preserving options for selective defences.

Tim Beardsley

Canadian budgets cut

Sacrificing all for space effort

Washington

CANADIAN science, already troubled by an uncertain financial future, took another body blow when the National Research Council (NRC) said recently that it would trim $20.5 million from its $400 million budget for the fiscal year starting next April.Particularly painful to those affected is that much of the money being trimmed from basic research efforts is being ploughed into the Canadian space programme, in particular for Canadian participation in the US space station.

The government makes no bones about the fact that some research projects will have to be curtailed or eliminated to support the space effort. In the next five years, NRC expects to realize an additional $125 million for its space station activities, but in the short term it has had to find $14.6 million for space activities from its existing budget.

Ralph Nicholls, a space scientist at York University, Ontario, and chairman of the Canadian advisory committee for scientific uses of the space station, says it is unfair to blame NRC for implementing government science policy. Nicholls says Canadian space scientists have also been hurt by the government's decision to eliminate the rocket and balloon programme, and now Canadian scientists find themselves in the same boat as their counterparts in the United States, who are having to wait for space shuttle flights to resume before their scientific projects can start again.

Although the cuts fell relatively evenly among the major divisions of NRC — biological sciences, chemistry, electrical engineering, astrophysics, construction research, mechanical engineering, aeronautics and physics — perhaps no area feels the cuts more keenly than radioastronomy. Canada has two major radioastronomy facilities; the Dominion Radio Astrophysical Observatory at Penticton, British Columbia, and the Algonquin Radio Observatory at Algonquin Park, Ontario. Both are under the control of the Herzberg Institute of Astrophysics, itself supported by NRC.

Institute director Donald Morton says there were plans to refurbish the Algonquin Park observatory to let it operate at shorter wavelengths. But when it became clear that the necessary $10 million would not be forthcoming, and the institute's budget was being trimmed by $1.1 million, attention focused on the viability of Algonquin Park. Although Morton insists no decision about the observatory's future will be made until next spring, he admits that prospects are not good.

Paul Wesson, a radioastronomer from the University of Waterloo, believes a decision to close the Algonquin Park Observatory would make Canada a "non-competitor" in international radioastronomy. Even without the refurbishment, the observatory is the largest steerable-dish antenna in North America devoted solely to radioastronomy. The planned improvement would allow the telescope to operate in the 1.5–4 mm range, making it the second largest millimetre-wavelength observatory in the world.

Wesson worries that because radioastronomy does not provide readily appreciated benefits to society, it is particularly vulnerable to future cutbacks.

Optical astronomy has also been hurt by the NRC budget adjustment. Sidney van den Bergh last month resigned as director of the Dominion Astrophysical Observatory in Victoria when the latest round of cuts was announced. He says the Victoria observatory has shut down one of its three telescopes, mothballed its optical shop and cut its support personnel to the absolute minimum.

Also struggling is TRIUMF (the Tri-University Meson Facility) in Vancouver, British Columbia. Last spring, NRC informed TRIUMF that its annual operating budget of $28.1 million would be cut by $4 million. Half that money was restored when it became apparent that TRIUMF would default on international contractual obligations without additional funds, but a $24 million budget is likely for the next fiscal year.

Earlier this year, the three Canadian granting councils announced a unique scheme to encourage industry contributions to scientific research by matching private funds with government money (see *Nature* **320**,102; 1986). Industry's willingness to participate in the scheme has yet to be demonstrated, however.

The Canadian Minister of State for Science and Technology, Frank Oberle, has promised to produce a national science and technology policy. With no firm timetable for its completion, many policy changes may be in place before the minister's plan is presented to parliament.

For those in government science posts both in and out of Ottawa, morale has plummeted. Many fear the current cuts are merely forerunners of larger cuts in the years to come, as the government shifts responsibility for basic research to universities and industry.

The furore created by NRC's cutbacks has prompted Oberle to form an independent panel that spokesman Dave Dale says will review the "whole overall macro picture" of NRC's role in Canadian science. Trying to sound optimistic, one NRC official said that although the council may be "more than slightly hurt, it is not mortally wounded".

Joseph Palca

New chairman for vice-chancellors

Sir Mark Richmond to CVCP.

PROFESSOR Sir Mark Richmond, vice-chancellor of the University of Manchester, will become chairman of the UK Committee of Vice-Chancellors and Principals in July 1987 in succession to Mr Maurice Shock, vice-chancellor of the University of Leicester. □

Accord renewed

Washington

AFTER a five-year hiatus, Canada and the Soviet Union have signed an agreement to resume mutual academic, scientific and cultural exchanges. Canada suspended official exchanges in 1981 following the Soviet intrusion into Afghanistan. The new agreement takes effect at the start of next year, and will run until the end of 1988.

The Canadian move comes on the heels of a similar exchange agreement signed this past summer by the United States and the Soviet Union (see *Nature* 322, 491; 1986). Anton Yurkowich of the Ministry of External Affairs says Canada was the last member of the North Atlantic Treaty Organization to renew basic exchange agreements with the Soviet Union among those that suspended them after the arrival of Soviet troops in Afghanistan.

Yurkowich says the decision to renew formal contacts is not tied to any specific political change, but rather that "the time was ripe to get back to talk again". The Ministry of External Affairs has obtained budget authority for up to 27 individuals for academic and scientific exchanges, and is negotiating with the Association of Universities and Colleges of Canada to administer the programme.

The major beneficiaries will probably be Canadian Slavonic scholars who have been denied access to Soviet library and archive material. Assured access to archival material was also keenly sought by the United States in its accord. Joseph Palca

Hubble Telescope taking shape

TECHNOLOGISTS from British Aerospace have just returned from California where they have been helping engineers from the European Space Agency's Technology Centre (ESTEC) to install the solar panels for the Hubble Space Telescope.

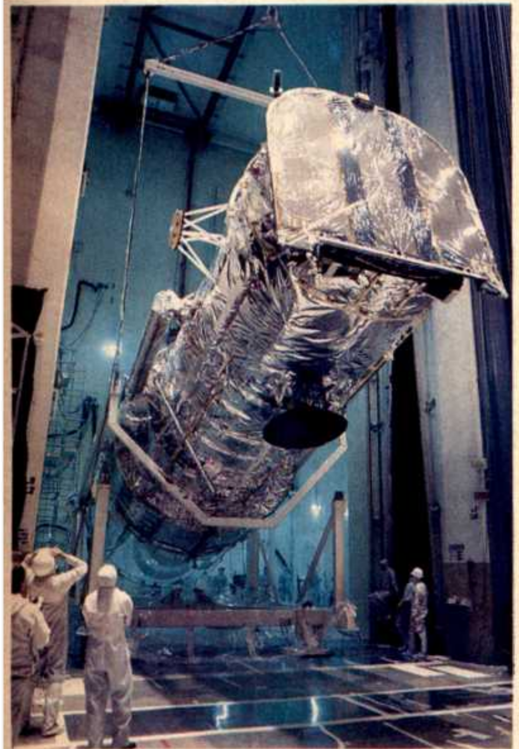
The arrays, built by a consortium led by British Aerospace, had been delivered in May to the telescope's construction site at the Lockheed Missile and Space Company complex in Sunnyvale. The arrays, the first to be designed to allow astronauts easily to retract and dismantle them, are contained on two panels unrolled from a single drum composed of more than 48,000 solar cells covering 356 square feet.

The retractability of the arrays and their ease of manipulation are two fundamental features of the design, as the telescope will have a 15-year life expectancy, three times that of the panels.

The telescope, due to be launched in two years time, will orbit at 593 km, allowing it to make observations free from Earth's atmosphere. It is designed to look seven times deeper into space, detect objects 50 times fainter and view them with 10 times greater clarity than the best telescopes on Earth. The Hubble Telescope will carry five scientific instruments that are replaceable and serviceable in orbit. They are the Faint Object Camera (FOC), the Faint Object Spectrograph, the High Resolution Spectrograph, the High Speed Photometer and the Wide Field/Planetary Camera.

Bill Johnstone

Hubble on test at Sunnyvale



Coral reefs

Okinawa runway under attack

Tokyo

THE battle to save one of the few thriving coral reefs left in Japan is about to reach its climax. The Okinawa prefectural government is expected soon to file a formal request with Japan's Construction Ministry to begin building an airport for jumbo jets on a reef of rare Pacific blue coral in Ishigaki Island, one of the southernmost islands of the Okinawa chain. Lined up to oppose the airport is an army of environmentalists from Japan and elsewhere.

The twin scourges of development and crown of thorns starfish (*Acanthaster planci*) have wreaked havoc with the coral reefs of the Okinawan Islands and the reefs of Ishigaki are rare remnants of what was once a thriving reef system.

Although Ishigaki already has an airport with a regular jet service to Okinawa, local government officials ten years ago proposed building a second airport with a longer 2.5-km runway to bring large jetloads of tourists direct from Osaka and Tokyo, and in 1979 the site of Shiraho lagoon was selected.

Shiraho was selected because a nearby hill (Karadake peak) could be razed and dumped in the lagoon to create a level area of 130 hectares, conveniently walled in by the outer reef. But the people of Shiraho town (population 2,000) are vehemently opposed to the plan.

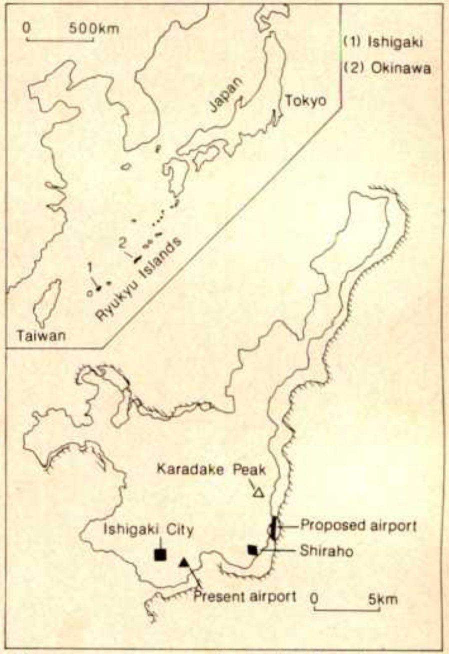
Two years ago, the Environment Agency planned a three-day survey of the reef. Local protesters tried to obstruct agency officials but were dissuaded by riot police and helicopters. Fearing further bad publicity, Ishigaki Town engaged Dr J. E. N. Vernon of the Australian Institute of Marine Science and Dr D. Kuhlmann of Humboldt University in Berlin to conduct a survey. But they were obstructed by Shoei Shirai, a local man appointed to help them, who later issued a report concluding that the reefs of north-west Ishigaki were the most valuable and the proposed site was not worth preserving. Vernon promptly disassociated himself from the report.

Vernon and other coral experts have been particularly struck by Shiraho's unusually extensive stands of Pacific blue coral (*Helipora cerulea*), a species listed under the Convention on International Trade in Endangered Species. Colonies tens of metres in diameter and 2-3 metres high stretch for hundreds of metres, constituting by far the biggest stands of this coral known in Japan.

In an attempt to allay concern, an informal advisory panel to the Governor of Okinawa that recommended construction of the airport has suggested that the coral be removed and 'replanted' elsewhere. But coral is not the only concern of opponents to the airport. Many have questioned the necessity of a second airport for such a small island, and there are fears that the 'spare' airport may be used for military purposes — Ishigaki is strategically placed at the southernmost tip of Japan.

Opposition to the airport is attracting worldwide attention. Last summer Jacques Cousteau wrote to the governor of Okinawa appealing for the plan to be dropped, and last month more than 500 environmentalists from Japan, the United States and the Philippines chartered a ship and called at Ishigaki to hold a protest rally on Shiraho beach.

In a ploy to delay construction, several hundred square metres of land needed for the airport apron that had slipped through the government buyer's hands were promptly bought up by the anti-airport group and sold off in doormat-sized lots to supporters. The government will now have to negotiate with about a thousand 'owners' to get the land back.



But the biggest obstacle for the airport planners is the fishing industry of Shiraho. Fishermen in Japan wield extraordinary power, and their consent must be won before construction projects affecting fishing grounds can go ahead.

Although the 600-member fishermen's cooperative of Ishigaki accepted Y500 million (£2 million) in 1980 as compensation for loss of fishing grounds at Shiraho, of which Y200 million was earmarked to pay off the cooperative's debts, the 30 fishermen of Shiraho have steadfastly refused to accept the agreement, and have taken their cooperative to court. As long as the disbursement of compensation remains blocked, airport construction cannot go ahead.

David Swinbanks

Yugoslavia

Serbian centenary "postponed"

SERBIA's Academy of Sciences and Arts, embroiled in a fierce political row with the authorities, has postponed the celebration of its centenary, due next week, until May. This decision follows fierce criticism by party and government leaders of an outspoken memorandum on the continuing social, economic and political crisis in Yugoslavia prepared by an academy group headed by Dr Antonije Isakovic, a vice-president of the academy. Under pressure, the academy has stopped work on the project.

Both the memorandum and the academy's decision to postpone its celebrations have been criticized. The Belgrade daily *Politika* on 23 October described the decision to postpone the celebrations as a kind of blackmail, and said that neither a handful of academics nor the president of the academy has the right to alter the date of the celebrations, which was "fixed by the history of its existence". *Politika* seems to have forgotten that, in 1974, the Soviet Academy of Sciences found it necessary to postpone the celebration of its 250th anniversary for a year.

The disputed memorandum deals with all aspects of Yugoslav life and presents a grim picture of bureaucracy, arbitrary decision-making, corruption, nepotism and a disregard for the law by those in positions of authority. Its advocacy of constitutional reform and the genuine participation of the public at large in decision-making has been attacked by the party theorists as anti-socialist, a betrayal of the legacy of Tito and evidence of a wish to break up the Yugoslav federation.

On education, though, the memorandum seems more concerned with keeping Yugoslavia together than with dismembering it. It castigates as divisive the present set-up in which each of the six republics and two autonomous provinces that constitute the federation has its own educational system (and academy of sciences). The memorandum says that countries without unified educational systems cannot expect to remain unified.

The memorandum also attacks a number of the educational concepts beloved of the party theorists, including that of "job-oriented" education, the result of which has been the over-production of ill-trained "specialists" unable to play a part in the social and economic life of Yugoslavia. It also says that the party campaign against elitism has meant a levelling down of standards and that the requirement that all persons employed in education should be "morally and politically suitable" has made the universities hot-beds of "careerism and conformity".

The Isakovic group also attacks the much-praised reform of the universities which hived off research into separate institutes, saying that the result has been the destruction of the laboratory base of the universities and the compartmentalization of science. The memorandum concludes that the universities therefore cannot play their essential part in the "technology-computer revolution". It advocates a thorough reorganization of scientific institutions in the interests of modernization, efficiency, greater financial investment and freer creativity.

Vera Rich

Ups and downs of French science budget

THE French research budget for 1987, up on the budget for 1986 as corrected after the present government came to power in April, but down compared with that planned for 1986 by the previous government, seemed to cause more excitement in the French National Assembly last week than any other aspect of next year's planned spending.

Leading socialist Jean-Pierre Chevènement was reportedly in fine form. His grand edifice of a ministry for research, technology and industry has been gradually dismantled, first by his own government (which trimmed off industry) and then by the present incumbents (technology has now gone, and the rump has been attached to the ministry of education). The ex-minister was exacting his revenge. He told the present science minister, Alain Devaquet, who is proud of his achievement in protecting the research councils and winning a modest rise in budgets against considerable internal opposition, that his budget was "faked and inflated". It was a "conjuring trick ... unworthy" of a sympathetic man like Devaquet.

The new minister rose to defend himself. The use of a single quantity, finance, was not the way to handle such a complex matter as French science, he said. Nevertheless, the national budget in government and industrial research and development would be up 8.5 per cent (on actual 1986 spending) to FF127,000 million (around £13,000 million), Devaquet claimed. A particularly big increase in defence research spending in the 1987 budget would offset decreases in support for industrial research in other areas. Criticized for this by socialists, Devaquet stoutly defended defence research: "I don't share the disdain of certain speakers for military research . . . It has done a great deal for many important industries."

However, André Giraud, the defence minister, will now have a research budget of FF30.8 million. Devaquet's budget for basic science is FF21 million. "So who is now the minister of research?" Chevènement taunted. Well, not M. Chevènement, anyhow . . .

Robert Walgate

X-ray astronomy

British–Soviet collaboration

A JOINT British–Soviet X-ray astronomy satellite may be launched in the early 1990s. Five British X-ray astronomy groups are now working out proposals for experiments to be submitted to the Soviet side at the end of December. If all goes well, the Soviet acceptance is likely to come by the middle of next year. The project, generated by a visit of a delegation from the British National Space Centre to the Soviet Institute of Space Research at the beginning of October, is viewed by both sides an important development. Dr Roald Sagdeev, director of the institute, said the meeting showed that Soviet and British scientists have good prospects for cooperation in space, especially in astrophysics, radioastronomy and the study of materials in space.

Until now, there has been relatively little direct cooperation between the two countries in space. The most significant project, so far, has been the British involvement in Roentgen, a module of four large X-ray telescopes to be mounted on the Soviet Mir orbital station next year.

But it was not the Soviet Union that initially proposed that Britain should take part but the Dutch, who wished to work in the hard X-ray/low-energy gamma-ray range using a coded-mask telescope, the principle of which was pioneered at the University of Birmingham. The first such telescope was flown on the Skylark rocket in 1976, and an improved version aboard Spacelab-2 in 1985. Professor Peter Willmore from Birmingham, who is now coordinating the proposals for the satellite, is reluctant to say too much about possible experiments. But the final proposal, he said, would be for a cluster of telescopes, using a variety of techniques to cover various energy bands.

The satellite, if all goes well, will be financed by each country paying for its own experiments and the Soviet Union covering the cost of the satellite and the launch. The project would therefore follow the normal financial pattern of Soviet collaborative projects, and has nothing in common with the purely commercial proposals, made by the Soviet Union in the wake of the Challenger disaster, that the Soviet Proton rocket could be used to launch satellites for foreign customers.

Vera Rich

Semiconductor rivalry

US bid to equal Japanese

Washington

THE beleaguered US semiconductor industry, hard pressed by superior manufacturing technologies from Japan, hopes to regain its competitive edge in world trade through a major cooperative manufacturing enterprise called Sematech. The Semiconductor Industry Association (SIA) will meet next week in Long Beach, California to give formal approval to the "manufacturing excellence project" that is already getting strong encouragement from US government.

Warren Davis of SIA says the industry's plight stems from the inability to develop effective manufacturing technologies for implementing innovative semiconductor designs. Davis believes no single US company is capable of taking on the scale of project needed for creating a new generation of manufacturing technologies.

Since early summer, representatives of the semiconductor industry have been holding discussions among themselves and with government officials on how best to proceed with Sematech. Some companies would like to see a full-scale manufacturing project, with the manufacturing muscle to beat the Japanese. Others would like a smaller pilot programme, with the technical fruits falling to individual companies. A compromise between the two approaches is likely.

Concern in Washington has been growing about the health of the US semiconductor industry, with several recent reports raising national security alarms over availability of semiconductor products (see *Nature* **323**, 483; 1986). Davis says if Sematech is to work, the government will have to give assurances that the effort will not be hindered by antitrust prosecution or unfavourable tax status. Referring to the Sematech project, Bruce Merrifield, assistant secretary of commerce for productivity, technology and innovation has spoken enthusiastically about the government's role as "the catalytic element of change", eliminating potential barriers to success. SIA officials were in Washington last week for talks with Merrifield and National Science Foundation (NSF) director Erich Bloch concerning Sematech. Although no government funds have been committed to the project as yet, commitments on taxes and antitrust issues could be worth millions of dollars. The Semiconductor Research Corporation, an industry supported consortium, recently received a grant worth $1.2 million from NSF and the Department of Defense, and could be a conduit for additional government funds.

Joseph Palca

Ozone depletion

US gets tough on CFC emissions

Washington

THE US government startled environmental conservatives and delighted radicals last week by calling for a complete halt to global emissions of halogenated alkanes thought to be destroying stratospheric ozone. The proposal was made in a paper on principles for an international protocol for protection of the stratospheric ozone layer issued by the State Department.

The United States will take the proposal as a starting position for the international negotiations beginning at Geneva on 1 December intended to put flesh on the bones of the Vienna Convention on Protection of the Ozone Layer. The convention has been signed by 25 countries since March 1985, but negotiations on a protocol to implement controls on chlorofluorocarbon (CFC) emissions have been postponed until the completion of informal technical meetings.

Anxiety about damage to the ozone layer, which filters out high-frequency ultraviolet radiation from the Sun, has grown in recent months. A 'hole' in the ozone layer appearing over Antarctica during the Antarctic winter (see *Nature* **321**, 729; 1986 and **315**, 207; 1985) is just now the subject of a study mounted by an inter-agency US team.

Although the cause of the 'hole' is not definitely known, the team's preliminary observations suggest that "a chemical mechanism is fundamentally responsible for the formation of the hole". Global ozone cover is thought to have decreased by 2–3 per cent in 1970–80, although measurements have been hard to interpret. CFCs also contribute to the greenhouse effect, by which the lower atmosphere may be warmed because of the effects of pollutants on the atmospheric radiation balance.

According to the State Department's proposal, a "long-term scheduled phase-out" of emissions should follow a near-term freeze on emissions of all fully halogenated alkanes. These include the widely used chlorofluorocarbons CFC 11, 12 and 113, believed to be the worst offenders, and halons 1211 and 1301. The proposal says that present uncertainties preclude the definition of a safe level of emissions, but that given the risks to health and the environment, the measures are justified. The proposed progressive reductions would be subject to change in the light of accumulating data.

The US proposal mentions in passing the need to develop incentives to persuade non-signatories of the convention to participate in the protocol.

The US industry group of chlorofluorocarbon manufacturers, the Alliance for Responsible CFC Policy, claims that the US proposal is not based on scientific principles. The alliance has proposed a cap on CFC production capacity and voluntary emission reduction measures to meet concern about the ozone layer.

The US proposal was based at least in part on a draft study by the Environmental Protection Agency, which estimates that incidence of skin cancer could increase significantly if emissions continue unchecked. But the industry alliance points out that the study has not yet been formally reviewed by the agency's scientific advisory board. A spokesman said that a phasing out of emissions did not necessarily mean an end to production, if some means could be devised to return CFCs to manufacturing plants for reuse.

Tim Beardsley

Deaths foretold

Washington

AN extra 40 million skin cancer cases and 800,000 deaths among people now alive and those born before 2074 are likely in the United States if CFCs are not further controlled, according to EPA's draft study* of stratospheric modification. The predictions stem from the expected increase in ultraviolet-B radiation reaching the Earth's surface as stratospheric ozone declines. The EPA study is largely responsible for the tough new US position on emission controls.

The predictions assume annual growth of 2.5 per cent for CFCs and 0.6 per cent for carbon dioxide, in line with current estimates. They also assume a continuation of present growth rates of nitrous oxide and methane (0.25 per cent and 1 per cent respectively). Equilibrium global temperature would under these conditions increase by between 3 and 9.5 degrees celsius by 2075, assuming the National Academy of Sciences' estimate of a 3 degree rise from a doubling of carbon dioxide.

EPA says that, as well as increasing skin cancers and cataracts, the increase of ultraviolet-B could promote the spread of herpes and leishmaniasis. There may also be effects on aquatic food chains and terrestrial plants, although information is scanty. But the study shows that estimates of risks are highly sensitive to CFC and halon growth rates; if the CFC growth rate is halved from 2.5 per cent to 1.2 per cent, estimated damage is reduced by 90 per cent. If a growth rate of 3.8 per cent is assumed, estimated damages increase by 400 per cent.

Tim Beardsley

An Assessment of the Risks of Stratospheric Modification (final draft), **submission to EPA Science Advisory Board, October 1986.**

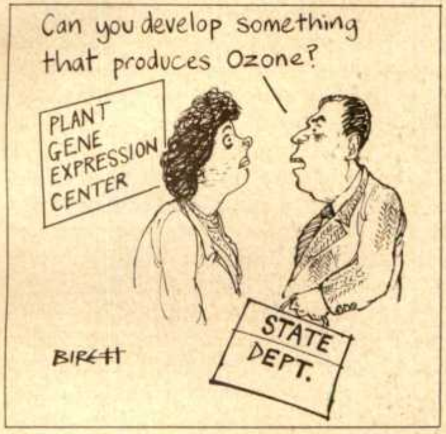
US agricultural research

Plant genes now fashionable

Washington

THE US Department of Agriculture (USDA), frequently criticized for its neglect of plant molecular biology, is now preparing to open a showcase institute with this theme at Albany, near San Francisco, California. The new Plant Gene Expression Center could herald a growth of interest from other government agencies.

At the outset, the new centre, under the aegis of USDA's Agricultural Research Service (ARS), will work on the transformation and regeneration of monocotyledons, the use of transposons to 'tag' and identify genes, molecular mechanisms of hormone action and environmental regulation of gene expression. USDA's annual support will be about $6 million, but extra funds will be sought from other sources.


Can you develop something that produces Ozone?
PLANT GENE EXPRESSION CENTER
STATE DEPT.
BIRETT

The scientific director will be Peter Quail, now at the University of Wisconsin. The centre will eventually house 10 research groups, some of whose leaders will also be professors at the nearby Berkeley campus of the University of California. Quail says the Berkeley link will be important; staff will be employees of the university, not civil servants, and visiting scientists will play a vital role. He hopes the centre will develop techniques for manipulating plant genes coding for informational molecules that will have long-term applications in crop improvement.

But while USDA is planning its own modest expansion of basic plant biology, the White House's Office of Science and Technology Policy is pushing for a more far-reaching increase in federal support. The view there (shared elsewhere) is that plant science has been eclipsed by the popularity of biomedicine, and urgent action is needed to redress the balance.

Indeed, the Office of Management and Budget hopes that there will be funds in the 1988 federal budget for an inter-agency initiative in basic plant biology . This proposal has been made by the Departments of Agriculture and Energy and the National Science Foundation and would support several new 'interdisciplinary research groups' in plant biology at universities, concentrating in five areas:

- microbial ecology and genetics in the rhizosphere
- chemistry of complex plant carbohydrates
- molecular biology and genetics of plant cells
- modelling of agro-ecosystems
- information technology.

Part of the impetus for the new initiative is the need to develop a 'predictive ecology' to help resolve the regulatory paralysis over environmental releases of genetically modified organisms. Whether those who hold the purse strings will support a major expansion remains to be seen.

Some academics are sceptical about the plan, saying that researchers can collaborate when they need to as things are and that the plan may simply be a way of dressing up a case for more money. Virginia Walbot of Stanford University believes that universities will find it hard to assemble true interdisciplinary groups without the promise of long-term support, which may be hard to get.

Peter Quail: aiming at better crops despite agricultural surpluses.

Others point out that the director of the Office of Science and Technology Policy, William Graham, whose support would be vital, has a background in defence, and may not be enthusiastic about biological research.

On a political level, the problems of US farmers are seen as stemming more from surpluses than from shortages, and biological research is seen — rightly or wrongly — principally as a way to make the surpluses bigger. **Tim Beardsley**

Grassroots unrest in plant biology

Washington

DESPITE the creation of the new Plant Gene Expression Center, seen by many as a belated attempt by (USDA) to jump on the plant biotechnology bandwagon, many academics believe research in basic plant biology has been seriously neglected.

The figures are telling; ARS's total budget will be $498 million this year, but that is all in-house and includes the cost of running laboratories. At least one-half is applied research such as plant breeding.

The Cooperative State Research Service, the USDA research arm which, jointly with state governments, supports the state experimental research stations, has a budget of $300 million, but only $40 million goes as competitive grants open to the university community. On one estimate, total federal support for innovative basic plant science open to competitive proposals from academics, including programmes of the Departments of Agriculture and Energy and the National Science Foundation, is only $110 million, compared with over $3,000 million spent on basic research by the National Institutes of Health in 1986.

ARS does support many small projects that use modern molecular biology techniques, but critics, such as Roger Salquis, president of Calgene Inc., and Stanford's Virginia Walbot agree it could have made better use of its opportunities. And the minuscule competitive grants programme, though increased in size in recent years. is not popular with academics; fewer than 10 per cent of proposals succeed, while grants (including overhead) vary between $50,000 and $80,000 in amount and last for only two years. Tim Beardsley

FRG–Soviet agreement

German deal put on ice

Hamburg

THE proposed agreement on research collaboration between West Germany and the Soviet Union seems to have been one of the first victims of the ill-feeling generated by Chancellor Helmut Kohl's controversial reference last month to Mr Mikhail Gorbachev, the general secretary of the Soviet Communist Party.

In an interview with the US news magazine *Newsweek*, Kohl had described Gorbachev as an "expert in public relations" and had likened him to Hitler's minister of propaganda, Josef Goebbels. Now Moscow has cancelled a visit during which Dr Heinz Riesenhüber, minister of research and technology in the federal government, was to have signed an agreement on the peaceful uses of nuclear energy and other matters reached only in July this year (*Nature* **322**, 488; 1986) after a decade of intermittent negotiation.

As things are, there is an exchange agreement between the Deutsche Forschung Gemeinschaft (DFG) and the Soviet Academy of Sciences but no agreement on scientific collaboration between the West German and Soviet governments. **Jürgen Neffe**

Remedy for British science

Sir—I agree with the spirit of your leading article (*Nature* **323**, 655: 1986) bemoaning the plight of British science, but I would question the political utility of the 'precepts' you urge the Advisory Board for the Research Councils (ABRC) to adopt to promote its case. In many ways they are all too familiar claims that those defending science have used in its defence. As such, they are as likely to be as unsuccessful in persuading Whitehall that its 'policies' — if that is not to dignify the disjointed initiatives of recent years too much — are damaging science as has been the more straightforward claim that 'we need more money'.

In many ways your precepts understate the real damage that is being done to the infrastructure of British science research. ABRC would be better advised, I believe, to make more of the wider national economic context in which science research and development finds itself, as well as to draw on what detailed information there is about the real damage that is being done to science at 'grass-roots level', the fundamental knowledge-generating culture of science.

On the first point, ABRC would do well to point out to its political masters that, according to the government's own Central Statistical Office, British industry, excluding oil, grew by only 1.5 per cent in 1979–85: little wonder that industry has had to look 'innovatory gift-horses in the mouth'. Moreover, where growth has occurred in the economy — in the financial sector — there appears to be little incentive on the part of financiers to invest in innovative research and development, basic, strategic or applied: money is best made out of money, on the financial markets.

At the more local level of laboratory life, my own research into the state of plant biotechnology and plant breeding in the United Kingdom reveals a failure by government to understand the research interests, abilities, needs and capacities for innovation of plant science. The speed and multidirectional nature of initiatives including privatization, rationalization and reorganization tend to pull (and literally so for its staff) science in opposite directions, breaking up the critical mass of research so vital to the future. Given the broader economic context sketched above, it is unlikely that scientific research and development is moving sideways into an expanding world of commercial opportunity. Management of science does need to be improved at national and local level: but clearly it can be effective only if it is informed by and sensitive to both.

A.J. Webster

Plant Biotechnology Project,
Cambridgeshire College of Arts and Technology,
East Road,
Cambridge CB1 1PT, UK

Pressure to publish

Sir—Writing in your New Journals Review issue, Mary Holmes admirably summed up the problems we face with the ever-increasing range of specialized scientific periodicals available (*Nature* **323**, 359–360; 1986). A cause contributing to this information explosion is the extraordinary pressure on scientists to publish. This is certainly noticeable in the medical field, where appointments and promotions frequently hinge on the sheer number of papers published in reputable-sounding journals. (A letter to *Nature* might get more brownie points than a paper in *The British Journal of Controversial and Spurious Knowledge* but 70 papers are more likely to get you a senior lecturership than one seminal paper.) This puts pressure on scientists to subdivide pieces of research to produce a number of papers rather than one complete publication.

The recent attempts in Britain to assess the research worth of entire departments or institutions on the strength of five publications has not proved particularly happy, but I do not see why the abilities of an individual should not be assessed on the strength of half-a-dozen papers. If enough institutions changed that part on the back of their job application forms which says "Publications" to "List not more than six of your publications" the pressure on scientists would change to trying to produce a small number of very good publications. If this happened, a lot of semi-vanity publications would cease, and much material currently appearing in glossy print would be reduced to a footnote saying something like "Further details of this experiment are obtainable from the author on request" with no great loss to the flow of scientific information.

Martin Guha

Institute of Psychiatry Library,
De Crespigny Park, Denmark Hill,
London SE5 8AF, UK

Scots elders

Sir—In the Opinion article "British science over the hill" (*Nature* **323**, 655; 1986), you say that three Scottish universities are more than 150 years old. The true figure is four: St Andrews (founded 1411), Glasgow (1451), Aberdeen (1495) and Edinburgh (1583).

A.M. Hetherington

113 Scotforth Road,
Lancaster LA1 4SD, UK

Isolated South Africa

Sir—*Nature* has a sound record (see **322**, 115; 1986) on human rights and freedom from repression for academics, but the conditions in South Africa would probably not have persisted if the grandees of the scientific community and the other learned academic communities had deployed their corporate influence against apartheid. In general, scientists and academics enjoy great prestige and privilege, and exert considerable influence within their own institutions. A large part of scientific research is supported by the public purse, either directly from government grants or from charitable foundations. It is not difficult for scientists to retreat into their laboratories until their research grants are up for renewal.

The cause of international scholarship is not more important than the fate and dignity of some 23 million black South Africans and five million white South Africans. One characteristic feature of the South African regime is its sensitivity to isolation. So long as South African athletes, diplomats, scholars and other professional and business people are granted free access to countries outside South Africa, especially the United Kingdom, Canada and the United States, the present regime in Pretoria knows that it will rule the Cape for a very long time.

The arguments against sanctions are based on previous experience when there was no common consensus for sanctions among the major European governments, Canada and the United States. There now seems to be a growing trend towards such a consensus.

J.A. Barnes

Department of Oral Biology,
University of Alberta, Edmonton,
Alberta, Canada T6G 2N8

Titles of references

Sir—I agree with John Maddox (*Nature* **322**, 681; 1986) that scientific literature is poorly written for the communication of the essential ideas of the research carried out.

One point he neglected to note is that a major contribution to this state is the obscure documentation of references. It is my opinion that a complete reference should always include the title of any cited work. I find this omission a great hindrance when researching related topics because I cannot determine beforehand whether an article is even remotely connected to my needs.

Unfortunately, I note that *Nature*'s reference style omits the titles of cited articles from journals. Your editors are contributing to this lack of communication as much as the authors.

Brian G. Latham

General Delivery, Yellowknife,
NWT, Canada X1A 2L8

How to tell when the sea rises?

Fears that, if the greenhouse effect is real, we shall drown rather than fry, require better ways of measuring sea level. There is a little progress to report.

THE present wave of interest in the linked problems of the atmospheric greenhouse effect and the integrity of the high-altitude ozone layer that spans the tropopause (see page 102, this issue) has also stimulated interest in what happens when the ice-sheets of Antarctica and Greenland begin to melt. Crude calculations of how mean sea level will be changed are easy enough. What we tend to overlook is that such a process would be novel only in degree, and that the consequences for the level of the oceans of the melting of the ice sheets of the last glaciation are still with us.

So much would have been plain to the participants in a conference on sea-level measurements and their interpretation held in Hawaii last year; some of the papers presented on that occasion have now been reprinted in the *Geophysical Journal* of the Royal Astronomical Society, and serve collectively as a reminder of the several influences that complicate the interpretation of the measurements, with all the complications arising from the need to distinguish mean sea level from the records of tidal gauges which are dominated by the variations due to solar, lunar and planetary tides.

It is remarkable, in the circumstances, that it has been possible to correlate changes of sea level along the south-eastern seaboard of the United States with the speed of the offshore Florida current, and mean sea level around the coast of Australia with a variety of oceanic and atmospheric phenomena, the Southern Oscillation linked with the El Niño phenomenon.

Obviously the time-span covered by a data set is crucial to the use that can be made of it. This is how it has been possible to infer, from tidal records around the margins of the North Sea, which parts of the coastline appear to be shrinking (the coast of England north of the Thames, for example) and which rising (the north coast of the Straits of Dover). But the statistical problems are formidable. People are looking for change of the order of 1 mm a year in data that may be distributed about the mean, after correcting for known seasonal variation, with a standard deviation of a few centimetres. Much of the incentive for the more detailed analysis of the effects of weather and ocean currents on sea level is the hope that it may, thereby, be possible to find meaningful signals in data sets that span shorter intervals. For what it is worth, those who believe that global warming may already be reflected in a steady rise of sea level suppose the magnitude of the change is approximately 2 mm a year.

That is why attention has also turned to the feasibility of direct synoptic measurements of movements of the Earth's surface topography of this order, where radio-interferometers on very long baselines (whence VLBI), already capable of measuring the distance across the Atlantic to within a couple of centimetres, are an obvious starting point.

W.E. Carter of the US Geodetic Survey (*Geophys. J. R. astr. Soc.* **87**, 3–13; 1986) seems to have explained to last years conference the accidental virtue of the US Defense Department's new navigation system (the Global Positioning System, which will have 18 satellites in three different orbital planes soon after the shuttle is in operation again) which has turned out to be a more convenient and cheaper version of VLBI. He offered the prospect that a sufficient density of suitable ground receivers, linked for accuracy with the VLBI network, should be able to monitor the level of tide gauges over reasonably short periods of time, less than a decade perhaps. But, rightly, people will first use this technique at the places where the surface level is changing much more rapidly; the post-glacial rebound of the area of the ex-Laurentian ice-shield (largely north of the Great Lakes) and Fennoscandia, where the change of level of the land surface approaches 10 mm a year (which is why the Great Lakes system is tilting towards the Mississippi Valley by rather more than 2 mm a year).

Another way of tackling the problem of direct measurement is to fix on some global quantity that can be measured accurately, and which depends somehow on the global sea level, so as to infer changes of the latter from measured changes of the former. What kind of quantity will fill that bill? The length of the day is obviously such a quantity, depending as it does on the moment of inertia of the Earth which in turn will depend on the mean sea level. The obvious difficulty is that there are many other influences on the length of the day, including both steady processes (such as the post-glacial rebound of North America) which progressively affect the Earth's moment of inertia, and dynamic processes (in the atmosphere or the oceans) which affect the Earth's rotation impulsively, and appear in the records as noise. How to distinguish all these factors?

Last year's symposium in Hawaii can have been left in no doubt of the difficulties after hearing the paper by W.R. Peltier, R.A. Drummond and A.M. Tushingham of the University of Toronto on the problem of post-glacial rebound (*ibid.* **87**, 79–116; 1986). Peltier is the one who, more than a decade ago, put the subject on a sound footing by showing that post-glacial rebound in North America and Scandinavia can be accounted for only by supposing that the present rate is controlled by viscous processes deep in the mantle of the Earth.

That is a curious tale. The Laurentian post-glacial rebound would not have taken 7,000 years if the compression caused by the loading of the ice had been a near-surface phenomenon. The calculations require that the effects of covering North America with ice would have been apparent 1,000 km or so beneath the surface, and that the rheology of material at that depth would imply a large viscosity. But in the past few years, it has emerged that that may be an illusion, one caused by what seems to be a phase transition at a depth of 640 km and which serves as a boundary through which influences cannot diffuse more quickly than the latent heat of the phase transition can escape.

For the once-glacial parts of the Earth's surface, the direction of sea-level change is determined by competition between isostatic rebound and the filling of the oceans with meltwater (which has long since ceased). In general, the changes involved are an order of magnitute greater than those now of interest — tens of metres per thousand years, not a millimetre or so each year. But much of the interest in Peltier's argument is that he is able to relate, for example, the changing topography of ice-sheets to the process of polar wander. Could it be that the best synoptic lookout for the predicted rise of sea-level after the greenhouse effect begins to bite will be a measurement such as that of the Earth's rotation poles?

Something of the kind is evidently necessary. The obvious indicator of global warming is the temperature of the Earth's surface, but that is even more complicated by noise (and incomplete coverage). But if the oceans act as a temporary sink for heat that would otherwise already be sufficient to make us over-warm, changing sea level might be the first indicator. **John Maddox**

Cytoskeleton

Dynamic microtubule dynamics

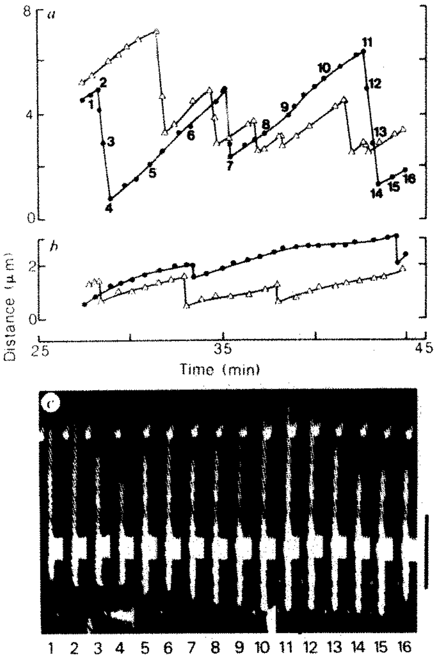
Robley C. Williams Jr, Michael Caplow and J. Richard McIntosh

WHEN assembly and disassembly of the microtubules from eukaryotic cells was first accomplished *in vitro* thirteen years ago, most observers thought the processes involved a simple reversible equilibrium between the polymeric microtubule and its tubulin subunits. Recent investigations reveal a livelier side of microtubule behaviour: an active growth and shrinkage of the polymer. A vivid example is the work of Horio and Hotani[1], who observed untreated microtubules *in vitro* in real time and found striking fluctuations in their lengths.

Microtubules in living cells come in all flavours of stability, from the ephemeral polymers of the mitotic asters to the sturdy stuff of which cilia and the protozoan cortex are constructed. Regulation of this diversity is, as yet, poorly understood, but it seems certain that the modes of microtubule assembly and disassembly lie at the heart of the matter. The assembly of tubulin subunits is accompanied by the hydrolysis of GTP. This reaction can produce a steady-state flux of subunits through the microtubule, a process termed treadmilling (see the discussion in *News and Views*[2]). Hydrolysis of bound GTP can, however, lag behind the incorporation of subunits[3]. The model that follows from this observation predicts the formation of inhomogeneous microtubules. In microtubules that are either growing or at a steady state of polymerization, the core of the microtubule is expected to consist almost completely of GDP-tubulin, whereas the recently added tubulin at the ends should form a GTP-tubulin cap[3,4]. Dissociation of a GDP-tubulin from the end of the exposed core is thought to be about 600 times faster than that of a GTP-bearing subunit from the cap[5]. Consequently, a microtubule is metastable: rapid disassembly results if the GTP-cap is destroyed (for example by dissociation of subunits from the end, by breakage, or by GTP-hydrolysis propagating endwards). According to this model, the macroscopic steady state of polymerization results from occasional rapid shrinkage of some microtubules, balanced by the slow growth of others.

This concept of microtubule behaviour was graphically supported by observations on the assembly of pure tubulin onto naturally occurring microtubule organizing centres[6]. Mitchison and Kirschner found that some of the polymers can shrink to disappearance while their immediate neighbours are growing. This phenomenon was presumed to occur via removal of subunits from, and addition to, the ends of the microtubule. Microtubules grown by elongation of preformed fragments also displayed a striking, simultaneous decrease in number and increase in the mean length of those remaining, both at steady state and on dilution. In the 'dynamic' instability model put forward to explain these observations, a given microtubule is in either a growing or a shrinking condition (phase), but is probably never static[7]. The relative rates of microtubule growth and GTP hydrolysis are thought to be the most significant factors in maintaining the GTP-cap[7]. (See ref. 2 for a discussion in *News and Views*.)

Is this model thermodynamically and kinetically sound? Is there perhaps some simpler alternative? The first question has been answered positively[8] in model studies that cast assembly and disassembly in terms of elementary reactions between soluble tubulin and plausible representations of the microtubule lattice. By choosing hypothetical values for the rate-constants, the model mimics the dynamic behaviour of microtubules in solution. The second question will not be as easily answered. Mitchison and Kirschner[6,7] explained the simultaneous increase in length and decrease in number of microtubules by an endwise, total disassembly of some polymers and addition of the released subunits to other, growing, microtubules. But immunogold labelling and electron microscopy studies[9] show that segments of preformed microtubules can also join together end-to-end to form a longer microtubule. This seemingly improbable but evidently quite efficient reaction, observed also with uncapped, taxol-stabilized, GTP-free microtubules (R.C.W. and M.C., unpublished observations), may contribute to the length redistributions on which the dynamic instability model is based.



Interconversions between the growing and shortening phases at the ends of a single microtubule. *a*, Length change at active end; *b*, change at inactive end; *c*, successive frame of a video image from one microtubule. (From ref. 1.)

Any doubt, however, that microtubules at steady state do fluctuate in length has now been dispelled by Horio and Hotani[1]. By means of real-time observation with dark-field optics, they can watch single microtubules change length (see figure). A given microtubule, when it is attached to the surface of a glass coverslip, can be followed through many cycles of shrinkage and growth. Their observations show directly that either end of each microtubule can exist in a growing or shrinking phase but that the length of a microtubule is never stationary. Micrometres of polymer, corresponding to thousands of subunits and perhaps 10–90 per cent of the length of a microtubule, assemble or disassemble between phase changes. The processes occurring at the two ends of one microtubule behave independently.

Not all these new results provide simple support for the picture developed by Mitchison and Kirschner[7,8], who explained the decrease in microtubule number and increase in microtubule length by asserting that the depolymerization phase would generally continue to completion, causing a microtubule to disappear from the population. Horio and Hotani's direct observation reveals no disappearing microtubules. Phase-reversal from shrinking to growth must be sufficiently probable, under the conditions of observation, to rescue essentially all microtubules from vanishing. One might think that this lack of conformity could be caused by the chamber used to make the observations. Attachment of microtubules to the glass surface might block their complete disassembly. But if this were the case, other microtubules not so close to the glass would disappear, thereby increasing the concentration of tubulin above that available at steady-state, so that the microtubules seen in the microscope would continually elongate. Because they neither disappear nor elongate, something other than simple dynamic instability must be going on.

Horio and Hotani also find, unexpectedly, that the end which undergoes more frequent reversals of phase, the 'active' end, grows and shrinks more rapidly than the infrequently reversing 'inactive' end. In current models, one imagines a GDP-bearing central region of the microtubule that gradually becomes longer as GTP hydrolysis proceeds on the incorporated subunits. At the active end, addition of GTP-bearing subunits should be more likely to outrun the spreading hydrolysis, leading to less frequent phase reversal than at the inactive end. It may be possible to reconcile the observed results with the GTP-cap model if GTP-hydrolysis occurs predominantly at the interface between the GTP-tubulin cap and the GDP-tubulin core (as suggested in ref. 10) and most rapidly at the active end.

Microtubule-associated proteins bind to the outside surface of microtubules, altering the energetics and kinetics of their assembly and disassembly. It seems likely that the GTP-cap persists in the presence of microtubule-associated proteins[5]. Yet when Horio and Hotani add these proteins to their microtubules, the frequency of phase reversal is greatly reduced. This result appears to be related to the observation[11] that microtubules containing microtubule-associated proteins do not show the length redistributions seen by Mitchison and Kirschner[6,7], but do undergo the treadmilling process. The function of microtubule-associated proteins in modulating the dynamic instability of microtubules thus may be physiologically important. This mechanism may, for instance, underlie the large, cell-cycle-dependent differences in lability of microtubules observed in photobleaching experiments with tubulin[12]. □

1. Horio, T. & Hotani, H. *Nature* **321**, 605 (1986).
2. McIntosh, J.R. *Nature* **312**, 196 (1984).
3. Carlier, M. -F. & Pantaloni, D. *Biochemistry* **20**, 1918 (1981).
4. Hill, T.L. & Kirschner, M.W. *Int. Rev. Cytol.* **78**, 1 (1982).
5. Carlier, M. -F., Hill, T.L. & Chen, Y. -D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 771–775 (1984).
6. Mitchison, T. & Kirschner, M. *Nature* **312**, 232 (1984).
7. Mitchison, T. & Kirschner, M. *Nature* **312**, 237 (1984).
8. Chen, Y. -D. & Hill, T.L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4127 (1985).
9. Rothwell, S.W., Grasser, W.A. & Murphy, D.B. *J. Cell Biol.* **102**, 619 (1986).
10. Caplow, M. & Reid, R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3267 (1985).
11. Farrell, K.W., Jordan, M.A., Miller, H.P. & Wilson, L. *J. Cell Biol.* (in the press).
12. Saxton, W.M. *et al. J. Cell Biol.* **99** 2715 (1984).

Robley C. Williams Jr is in the Department of Molecular Biology, Vanderbilt University, Nashville, Tennessee 37235 USA; Michael Caplow is in the Department of Biochemistry, University of North Carolina, Chapel Hill, North Carolina 27514, USA; J Richard McIntosh is in the Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309, USA.

Natural selection

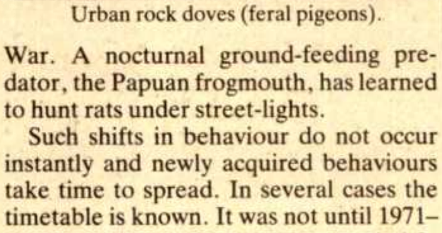
Rapid evolution of urban birds

Jared M. Diamond

A MAJOR dispute in evolutionary biology concerns the rate at which species evolve to occupy new niches. How suddenly did *Homo sapiens* replace *H. erectus*, and how soon after the demise of dinosaurs did large mammals arise? It is inevitably difficult to answer such questions using fossil evidence. Yet we live among many little-studied cases of rapid evolution, as species occupy habitats created by man. An example is the very recent formation of an urban avifauna in Papua New Guinea, summarized in a posthumously published paper by Harry Bell (*Proc. Western Fdn Vert. Zool.* **3**, 1; 1986).

In Europe, where towns have existed for thousands of years, many species of birds have become human commensals. Rock doves, starlings and house sparrows thrive in concrete habitats, and the latter species is now so closely associated with man that its original niche is unknown. In Papua New Guinea, however, the first town was founded a mere 112 years ago, and today only 16 towns have human populations exceeding 4,000. Typical town habitats are very unlike the New Guinea rainforest and consist of scattered trees and ornamental shrubs, mostly of introduced species. How quickly are bird species adapting to these new habitats, and what makes for a successful colonist?

In some cases the pre-adaptations are clear. The greater wood-swallow, which traditionally sallies from tall ridge-top trees over the primary rainforest, in towns sallies from radio towers and telephone poles. Shining starlings transferred their nests from holes in trees to vents of air conditioners, thereby becoming a pest. Swiftlets that would otherwise nest in caves now use tunnels excavated by the Japanese army during the Second World War. A nocturnal ground-feeding predator, the Papuan frogmouth, has learned to hunt rats under street-lights.


Frank Lane.

Urban rock doves (feral pigeons).

Such shifts in behaviour do not occur instantly and newly acquired behaviours take time to spread. In several cases the timetable is known. It was not until 1971–74 that a flock of black kites settled in the town of Lae to feed on road kills, although the town was founded in the 1930s. Only since 1983 have brahminy kites learned to feed on road-killed toads in Port Moresby. Goldie's lorikeet, which was formerly a rare species occurring in primary forest, arrived in the 1970s to feed on casuarina seeds in highland towns, where it is now the most abundant bird. Lemon-bellied flycatchers expanded from the savanna into the town of Wau between 1976 and 1978, 50 years after Wau was founded.

Checkered distributional patterns resulting from interspecific competition are common among closely related New Guinea bird species (my own work, in *Ecology and Evolution of Communities* eds Cody, M. & Diamond, J.M. p. 342; 1975). Bell's analysis of the town distribution of fruit doves (genus *Ptilinopus*) gives the first indication of how quickly the checkerboards develop. Three lowland species (orange-bellied, orange-fronted and pink-spotted) are each now resident in one or more towns, but no town except Popondetta supports more than one of the species, so that the distributions form a checkerboard. In natural habitats the orange-bellied and pink-spotted fruit doves are geographically widespread, the pink-spotted extends to higher elevations, but the orange-fronted is absent from the region that includes the towns of Lae and Madang. Town distributions are as follows: the orange-fronted is common in all lowland towns throughout its geographical range; the orange-bellied is found as a town bird only in Madang and Lae (and in Popondetta where it is less common than the orange-fronted and is possibly declining); and the pink-spotted occupies only Bulolo, above the altitudinal ceiling of the orange-fronted and orange-bellied. These results suggest that the orange-fronted has a competitive advantage in towns, that the orange-bellied succeeds only in the absence of the orange-fronted and that the pink-spotted succeeds only in the absence of both relatives. If so, this distribution must have arisen within the past 40 years and may not yet have reached equilibrium in Popondetta. There is a similar distribution in New Zealand, where two introduced starlings (the European starling and the Indian myna) have both acquired the habit of surveying pastures for food while perching on the back of a sheep, but in any given pasture only one of the two species occupies the sheep.

The successful town colonists are unrepresentative of the New Guinea avifauna as a whole in their preferences among natural habitats. Of species present in natural habitats near towns, only 4 per cent of the rainforest species but 33 per cent of the savanna species and 50 per cent of the second-growth species occupy towns. These differences partly reflect the varying similarities between the natural habitats and town habitats (least like rainforest), partly the behavioural plasticity of second-growth species. The very low success of rainforest species has grim implications, because it means that spreading urbanization will threaten New Guinea's famous rainforest avifauna.

As Bell notes, New Guinea town avifaunas are still far from the equilibrium that now exists in Europe. Changes in behaviour and status are occurring rapidly. The species/abundance relation remains unusually equitable: many species are vying for space in New Guinea towns, in contrast to the pattern of a few dominant species in older towns in other countries. This panorama of biology in rapid flux provides opportunities for field biologists to address some basic questions. Does the sudden acquisition of new behaviour depend on a genetic change or on culturally transmitted learning? How quickly does morphological and physiological evolution occur among colonists of the new habitats? We know that house sparrows living in wet or dry and hot or cold areas of North America differed morphologically when studied a century after their arrival (Johnston, R.F. & Selanders, R.K. *Science* **144**, 548; 1964) and that the advance of house sparrows in tropical Australia halted for 60 years before their final surge across the Tropic of Capricorn. Biologists neglected those and innumerable other opportunities to study natural selection in action. Such opportunites still exist in New Guinea and elsewhere. ☐

Jared M. Diamond is Professor of Physiology at the University of California Medical Schoool, Los Angeles, California 90024, USA.

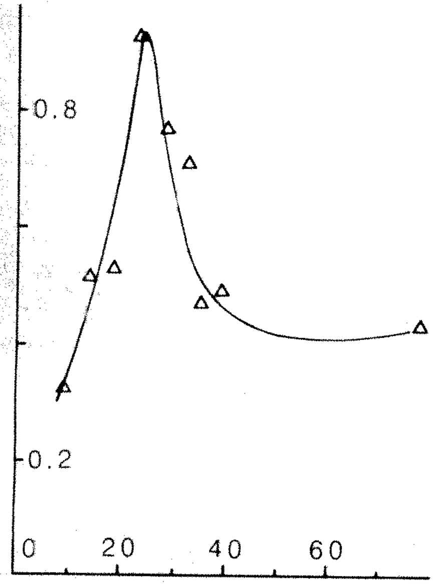
Metallic superlattices

The supermodulus syndrome

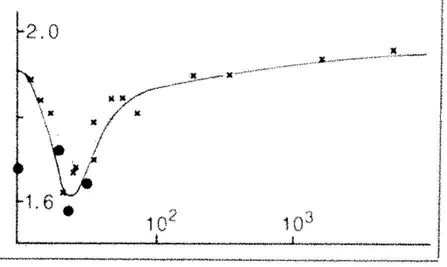
Robert W. Cahn

SYNTHETIC compositionally modulated epitaxial semiconductor structures, or semiconductor superlattices, were first proposed by Esaki and Tsu in 1969. Because of their remarkable electronic and optical properties, these structures have been investigated intensively since then. The metallic equivalent of semiconductor superlattices, synthetic modulated metal films, began to excite interest in 1977, but until now studies on these films have been scientific rather than technological. Recent work has succeeded in using synthetic, bimetallic modulated films for the spectral analysis of X rays[1], the investigation of extremely low diffusivities of one metal in another[2,3] and the study of phase transformations[4].

In 1977, Yang *et al.*[5] prepared a series of multilayer films of the metal pairs Au–Ni and Cu–Pd. The elastic stiffness of the films was measured by a primitive method that allowed the film to bulge under gas pressure over a small hole; the bulge height is determined as a function of gas pressure. The stiffness thus measured is a biaxial modulus. For Au–Ni the stiffness doubles for a critical thickness near 2 nm; for Cu–Pd, the effect is less pronounced. Recently, Baral *et al.*[6] used a more conventional, vibrating-reed approach to measure Young's modulus. The figures show their results[7] for Cu–Ni of constant 50/50 Cu/Ni ratio but varying periodicity and, for comparison, results with measurements of surface-wave velocity for 50/50 Cu–Nb films. The elastic anomaly near 2-nm thickness can evidently be a maximum or a minimum, although taken together, all published measurements[8] indicate that the anomaly is usually an enhancement and only rarely (to adopt the latest neologism) a dehancement. In Cu–Au films, however, there is no change. In some of the systems that show a modulus peak at a critical periodicity, there is also a peak of critical stress for plastic deformation.



Left, Variation of flexural modulus (in TPa, vertical scale) with wavelength (in Å, horizontal scale) for Cu–Ni. Below, Surface-wave velocity ($\times$ 10^5 cm s^{-1}, vertical scale) of Nb–Cu superlattices as a function of bilayer thickness (in Å, horizontal scale).



There are two main rival theories that seek to interpret this 'supermodulus' effect. Clapp[9] proposed a purely electronic interpretation in which the compositional modulation creates additional Brillouin zones which can touch the Fermi surface and thereby stabilize the modulated structure. If this happens, then an elastic strain will tend to drag the zone and Fermi surface out of contact, which acts in the direction of structural destabilization and thereby enhances the stiffness. (The converse effect is also conceivable.) As various theorists have pointed out (for example, ref. 10), this hypothesis predicts several critical periodicities that should lead to anomalously enhanced (or dehanced) stiffness.

The alternative view[11] is that the supermodulus effect is caused by coherency strains at the metal/metal interfaces, which are postulated to displace atoms so far from their normal sites that they are no longer within the hookean parts of their interaction potential. This approach can also interpret both enhancements and dehancements, but cannot explain the origin of the sharpness of the maximum so easily. For large periodicities, interface dislocations tend to form and release the coherency strains, but this still leaves unexplained the behaviour at periodicities of less than 2 nm. One interesting possibility is that the dehancement of stiffness in Cu–Nb films might result from glass formation by interdiffusion, the so-called solid-state amorphization reaction[12]. Cu–Nb would be a reasonable candidate system for this process, though it has not been investigated with the amorphization reaction specifically in view. It is interesting that Lin and Spaepen[13] find that in Ni–Nb multilayer films, metallic glass is formed spontaneously during film deposition (that is, without anneal) for layer thicknesses smaller than 3 nm.

Very recent work using Cu–(Ni,Fe) multilayers[14] shows the existence of a minimum grain size (~0.1 μm) at 2.6-nm periodicity. At this periodicity, the structure is seen to be defect-free; at smaller periodicities, partial interdiffusion reduces the sharpness of the interfaces and allows grains to coarsen; at periodicities greater than 2.6 nm, interfacial dislocations reduce the elastic coherency strains. These mechanisms can account for the modulus maximum at 2.6 nm. In other recent work[15], elastic moduli are calculated on the basis of the coherency strain model for Cu–Au and Cu–Ni multilayers as a function of periodicity distance, and roughly agree with experimental values

The debate between the electronic and coherency strain hypotheses (see refs 8,11) remains wide open, but it does increasingly look as though the coherency strain model will turn out to be correct

No practical use has been made yet of supermodulus enhancement, but the situation may be about to change Stiffness is a major concern in the development of improved aircraft alloys, and Al–Li and Al–Li–Be alloys, which are stiff and light, are now receiving concentrated attention A very new approach is to make massive aluminium alloy pieces, large enough to machine wing spars, by sequential evaporation[16] such alloys have substantially enhanced flow stress as well as stiffness, but there has been no report of a supermodulus effect at critical periodicities In any case, the periodicities used were 30–40 times larger than in the thin-film work, where peaks in flow stress have been observed at $\sim$2 nm periodicities[15] It remains to be seen whether alloy pieces of useful size can be made with perodicities around 2 nm, and whether an aluminium-based system will show the supermodulus effect Yahalom[17] has just reported a very promising new approach to making nanometre metallic multilayers by pulsed electrodeposition, applied initially to the Cu–Ni system He specifically suggests applying his method to the production of springs □

1 Underwood, J H & Barbee, T W *Nature* **194,** 492 (1981)
2 Greer, A L *Scripta Met* **20,** 457 (1986)
3 Cammarata, R C & Greer, A L *J Non-Cyst Solids* **61** and **62,** 889 (1984)
4 Tsakalakos, T *Scripta Met* **20,** 471 (1986)
5 Yang, W M C, Tsakalakos, T & Hilliard, J E *J appl Phys* **48,** 876 (1977)
6 Baral, D, Ketterson, J B & Hilliard, J E *J appl Phys* **57,** 1076 (1985)
7 Kueny, A *et al Phys Rev Lett* **48,** 166 (1982)
8 Cammarata, R C *Scripta Met* **20,** 479 (1986)
9 Clapp, P C in *Modulated Structure Materials* (ed Tsakalakos, T) 465 (Nijhoff, Dordrecht, 1984)
10 Falco, C M & Schuller, I K in *Synthetic Modulated Structures,* (eds Chang, L L & Giessen, B C) 339 (Academic, London, 1985)
11 Tsakalakos, T & Jankowski, A F *A Rev Mat Sci* **16,** 293 (1986)
12 Schwarz, R B & Johnson, W L *Phys Rev Lett* **51,** 415 (1985)
13 Lin, C J & Spaepen, F *Acta Metall* **34,** 1367 (1986)
14 Jankowski, A F & Shewbridge, J F *Mat Lett* **4,** 313 (1986)
15 Imafuku, M, Sasajima, Y, Yamamoto, R & Doyama, M *J Phys* **F16,** 823 (1986)
16 Bickerdike, R L *et al Int J Rapid Solidification* **1,** 305 (1986)
17 Yahalom, J in *Materials and Processing Report* Vol 1 No 7 (October 1986)

Robert W Cahn is in the Department of Metallurgy and Materials Science, University of Cambridge, Cambridge CB2 3QZ, UK

Molecular evolution

Deciphering divergent codes

L A Grivell

To evolutionary biologists the protists must appear a motley collection of organisms with diverse characters, defying simple classification, and with many unusual features whose original selective advantage can sometimes only be guessed at The discovery last year[1-5] that ciliate protozoa make use of an unusual version of the genetic code, in which UAA and UAG codons specify glutamine instead of stop is entirely in character and it offers molecular biologists a unique opportunity to piece together events in early evolution that may have led to the appearance of the present-day genetic code Clues to what may have happened are presented in a recent article by Hanyu *et al* [6]

Until 1981, the genetic code was believed to be the same in all organisms and it was therefore quite reasonable to assume that it had been fixed before the divergence of all life forms, any changes being impossible because they would make nonsense of messages for existing proteins This belief was shaken by the finding[7] that different versions of the code are used in the mitochondria of different organisms and that some of these changes must be of relatively recent origin

The simplest explanation for this extraordinary degree of genetic tolerance within mitochondria seems to be that only a limited set of proteins are encoded that function entirely within the hydrophobic environment of the inner mitochondrial membrane Mutations leading to changes in codon assignment, which would be lethal in a more complex genetic system with a wider range of gene products, are apparently acceptable This may also hold for the simple prokaryote *Mycoplasma capricolum,* in which, as in many mitochondria, the stop codon UGA is used to specify tryptophan (see ref 5 for a review)

How does the deviant code in the ciliates fit this picture? Are these organisms like mitochondria in being able to absorb changes in their genetic dictionaries, or are we looking at a relic of a primitive genetic code, one which possibly preceded the emergence of the so-called universal code? The first alternative seems unlikely in the light of the complex and definitely eukaryotic characters of the ciliates The second, in its simplest form, would seem to require that all bacteria, plants, fungi and animals have independently but consistently switched UAA and UAG codons to signal stop Hanyu *et al* [6] suggest a possible solution to this problem

These authors show the occurrence in *Tetrahymena* of three different transfer RNAs that can attach glutamine The first has the anticodon UmUG (m indicating methylation), capable of decoding the two normal glutamine codons CAA and CAG and is probably the counterpart of the glutamine-specific tRNA present in other organisms The others are unusual and consist of a $\text{tRNA}^{\text{Gln}}_{\text{UmUA}}$ and a $\text{tRNA}^{\text{Gln}}_{\text{CUA}}$ The first of these recognizes both UAA and UAG in natural messenger RNAs in an *in vitro* protein synthesizing system, while the second recognizes UAG only Sequence comparison shows that all three tRNAs are related, with $\text{tRNA}^{\text{Gln}}_{\text{UmUA}}$ and $\text{tRNA}^{\text{Gln}}_{\text{UmUG}}$ showing 81 per cent homology and $\text{tRNA}^{\text{Gln}}_{\text{UmUA}}$ and $\text{tRNA}^{\text{Gln}}_{\text{CUA}}$ 95 per cent homology Hanyu *et al* therefore conclude that the two unusual glutamine-specific tRNAs evolved from the normal one early in ciliate evolution by gene duplication and divergence, with the UAG-specific $\text{tRNA}^{\text{Gln}}_{\text{CUA}}$ derived from the $\text{tRNA}^{\text{Gln}}_{\text{UmUA}}$ to give a strengthened response to UAG

They go on to argue that because, apart from mitochondria and *Mycoplasma,* all eubacteria and eukaryotes normally use UAA, UAG and UGA as termination codons, it is more likely that two of these codons failed to specify termination during the evolution of the ciliates, rather than being acquired as stop codons by the other organisms in a process of convergent evolution They therefore propose that although UAA and UAG were probably stop codons in the archetypal ciliate, they must have been rarely used and that weak suppressor tRNAs for these codons may have existed at a very early stage Mutations creating UAA and UAG codons within protein-coding genes would not have been lethal for such prociliates Any subsequent mutation which allowed acquisition of an anticodon capable of perfect recognition of these codons would then result in the fixation of their use as amino-acid codons A G-to-A transition in the anticodon of a duplicated glutamine-specific tRNA gene is one of the simplest ways in which this could have occurred

This picture has some attractive aspects First, it is consistent with the suggestion by Woese[8] that the primitive genetic code may have been highly ambiguous and error-prone and that even though considerably fewer than the present 20 amino acids may have been encoded, it would have been important to minimize the number of codons not specifying an amino acid Second, it is consistent with molecular phylogenetic data[4,9] indicating that the ciliates branched off from the primitive eukaryotic ancestor very early in evolution, quite possibly at a time when protein translation was still tolerant of changes in the genetic code Finally, it is consistent with more recent patterns of change in evolution, in which gene duplication followed by the acquisition of a new function by one of the duplicated genes is common Despite these attractive features, however, the process proposed by Hanyu *et al* [5] tends to underplay the fact that many other organisms are able to recognize and translate termination codons which are used either to modulate the synthesis of certain proteins, or to allow the insertion of modified amino acids (for reviews, see refs 10,11) Recognition of termination codons *per se* is thus not unusual How they are recognized depends on subtleties of tRNA structure, modification or codon context The usage of UAA and UAG as normal codons in the ciliates may be an extreme form of this recognition process, perhaps one that has developed secondarily as part of the peculiar life style of these organisms

If, on the other hand, the existence of tRNAs capable of the recognition of stop codons is the relic of an intrinsic property of a primitive and error-prone code, additional variations in the code in other evolutionary groups might be expected, particularly in those subject to similar evolutionary pressures and who branched off the main eukaryotic line either before or around the same time as the ciliates Based on sequence comparisons of ribosomal RNA (ref 9), *Dictyostelium, Euglena* and the *Trypanosoma* represent even earlier divergence points from the mainstream of eukaryotic descent But neither *Dictyostelium* nor *T brucei* shows any deviation from the standard code Data on *Euglena* and other ancient groups, such as the dinoflagellates, the red algae and *Pelomyxa,* are eagerly awaited

Finally, a couple of other intriguing questions remain The first relates to the mechanism by which protein synthesis is terminated As Hanyu *et al* [6] point out, higher eukaryotes seem to use only one release factor for all three termination codons It will therefore be of interest to see how the *Tetrahymena* factor, which should recognize only UGA, differs from its known eukaryotic counterparts The other question is whether any organism exists or has existed in which stop codons are not used at all, so that translation is ended by 'run-off' synthesis It may be worthwhile to delve further among the protists, the answer could be in there somewhere □

1 Caron, F & Meyer E *Nature* **314**, 185–188 (1985)
2 Preer J R Jr Preer, L B Rudman, B M & Barnett A J *Nature* **314** 188–190 (1985)
3 Helftenbein, E *Nucleic Acids Res* **13**, 415–432 (1985)
4 Horowitz, S & Gorovsky M A *Proc natn Acad Sci U S A* **82**, 2452–2455 (1985)
5 Fox T D *Nature* **314**, 132–133 (1985)
6 Hanyu N Kuchino, Y , Nishimura, S & Beier, H *EMBO J* **5** 1307–1311 (1986)
7 Grivell, L A (1983) in *Mitochondria 1983 Nucleo-Mitochondrial Interctions* (eds Schweyen, R J , Wolf, K & Kaudewitz F) 25–45 (De Gruyter Berlin, 1983)
8 Woese C R *Proc natn Acad Sci U S A* **54**, 1546–1552 (1965)
9 Sogin, M L Elwood H J & Gunderson, J H *Proc natn Acad Sci U S A* **83**, 1383–1387 (1986)
10 Hatfield D *Trends biochem Sci* **10**, 201–204 (1985)
11 Chambers I *et al EMBO J* **5**, 1221–1227 (1986)

L A Grivell is in the Section for Molecular Biology in the Laboratory of Biochemistry, University of Amsterdam, Kruislaan 318, 1098 SM, Amsterdam, The Netherlands

---

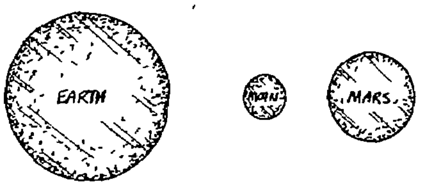
Cosmogony

Moon's birth shocked Earth?

Alan P Boss

THE most promising explanation of the birth of the Moon is the impact of a Mars-sized body with the Earth during the final phase of planetary accumulation If such an impact occurred nearly tangentially, the high angular momentum of the Earth–Moon system can be understood, one of the major attractions of this theory The first detailed computer simulations of such a giant impact have now been completed[1] and can be compared with lunar and geochemical data[2] to help isolate the manner in which the Moon may have formed

Thinking on the origin of the Moon has undergone a remarkable transformation in the past few years[3] Recent work on the three classical hypotheses of fission, capture and binary accretion has uncovered flaws in their mechanics serious enough to warrant their rejection But at the same time a relatively new hypothesis has become (not entirely coincidentally) quite popular This hypothesis invokes the collision of a Mars-sized or larger protoplanet with the growing protoearth[4,5], followed by formation of the Moon in Earth orbit from the debris of the explosion



Collision at 10 km s^{-1} of a Mars-sized body with the protoearth would produce a colossal explosion that may have led to lunar formation

When originally proposed a decade ago, the giant impact hypothesis received little attention, largely because giant impacts were thought to be highly improbable But a new variant of the general theory of terrestrial planet formation predicts that the final phases of formation primarily involved the accumulation of bodies ranging from about the size of the Moon (around one-hundredth the mass of the Earth) to three times the mass of Mars (three-tenths of the Earth)[6] The appearance of a general theory which predicts the occurrence of Mars-sized impacts on a protoearth of around the present Earth mass with essentially unit probability, coupled with the apparent demise of the competing theories, has encouraged serious efforts at exploring the consequences of the giant impact hypothesis

Benz and his colleagues[1] use a computer code to model the detailed physics of giant impacts much more accurately than previous efforts[7] The outcome of a giant impact depends sensitively on impact parameter (the offset between the centres of the colliding bodies) and relative velocity Small impact parameters require high relative velocities to achieve the angular momentum of the Earth–Moon system Benz *et al* found that small impact parameters lead to complete disruption of the impactor and to accumulation of this impact ejecta by the protoearth Their favoured model involves a large impact parameter where the impactor grazes the surface of the protoearth, survives the collision largely undispersed and quickly reforms into a cohesive body The protomoon is trapped in Earth orbit by a combination of dissipation of relative velocity during the collision, acceleration by gas pressure gradients in vaporized phases, and gravitational torques with the deformed protoearth Because the orbit of the protomoon initially has a perigee within the Roche limit of the Earth (around three Earth radii), the protomoon may eventually be tidally disrupted, forming a de-volatilized disk of matter which may accumulate into the final Moon More recent calculations by Benz *et al* include iron cores in both bodies, and show that the iron core of the impactor can be swallowed by the protoearth, leaving the impactor mantle to form the protomoon

These models illustrate the main successes of the hypothesis giant impacts provide a reasonable mechanism for injecting roughly a lunar mass of matter into Earth orbit, with the right amount of angular momentum, and with the matter depleted in iron and volatiles, as required by chemical analysis of lunar samples Furthermore, an origin as violent as proposed seems consistent with the large-scale melting required by lunar magma ocean models

One difficulty with this version of the giant impact hypothesis is now apparent Ringwood[2] has re-emphasized the strong similarities in abundance of siderophile elements in the Earth and the Moon The models of Benz *et al* imply that the protomoon formed almost entirely out of the impacting body, thereby requiring the impactor, formed elsewhere in the terrestrial planet zone[6], to have a mantle very close in composition to that of the Earth This improbable requirement spoils the beauty of the hypothesis But there may be a way out

Melosh and his colleagues[8,9] have also produced detailed models of giant impacts Contrary to Benz *et al* , they find that large-scale vaporization accounts for the formation of a prelunar disk about the Earth The discrepancy appears to be caused by different assumptions about the equation of state and by difficult numerical techniques, with the work by Melosh *et al* probably coming closer to reality Because vaporization does not discriminate between impactor and impactee, Melosh's prelunar disk is a roughly equal mixture of impactor and protoearth Making the Moon out of a mixture of impactor and protoearth mantle seems to solve the siderophile problem, because the protoearth composition would also be altered by the addition of the bulk of the impactor, with the final Earth mantle being composed of as little as about 70 per cent protoearth mantle (assuming a three-tenths Earth-mass impactor striking a seven-tenths Earth-mass protoearth), a proportion closer to that of Melosh's prelunar disk

Clearly many details remain to be investigated and other problems, such as giant impacts totally melting the Earth and producing a massive crust incompatible with the present Earth, will require explanation If divining the origin of the Moon were a paying proposition, venture capital firms would certainly be well advised to invest in improved models of giant impacts ☐

1 Benz, W *et al* *Icarus* **66**, 515 (1986)
2 Ringwood, A E *Nature* **322**, 323 (1986)
3 Boss, A P *Science* **231**, 341 (1986)
4 Hartmann W K & Davis D R *Icarus* **24**, 504 (1975)
5 Cameron A G W & Ward, W R *Lunar Science Conf VII*, 120 (1976)
6 Wetherill, G W *Science* **228**, 877 (1985)
7 Cameron, A G W *Icarus* **62**, 319 (1985)
8 Melosh H J & Sonett, C P in *Origin of the Moon*, 621 (Lunar Planet Inst 1986)
9 Kipp M E & Melosh H J *Lunar planet Science Conf XVII*, 420 (1986)

Alan P Boss is in the Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road NW, Washington DC 20015, USA

Fusion energy

Highest temperatures yet

R S Pease

DEUTERIUM gas in the Tokamak Fusion Test Reactor at Princeton was heated this summer to a temperature of 230 × 10^6 K (20 keV), perhaps the highest temperature yet recorded in a laboratory[1,2] It lies close to the temperatures planned for fusion reactors, and paves the way for the planned production of megawatts of fusion power in deuterium−tritium experiments in 1989 These promising results bring us closer to the goal of controlled nuclear fusion energy

What does this experiment amount to, and how does it relate to other attempts? The Tokamak Fusion Test Reactor experimental assembly is a toroidal electric discharge in a strong longitudial magnetic field Three other large tokamaks are now operating the Joint European Torus at Culham, JT-60 at Tokai in Japan and D III-D at San Diego In the Tokamak Fusion Test Reactor, the torus major radius is 2 4 m and its minor radius is about 1 m The discharge current in the high-temperature experiments is about 1 MA and lasts about 5 s, this current provides the magnetic field which confines the high-temperature deuterium plasma and provides the necessary thermal insulation, it also provides a relatively small amount of heating which on its own heats the plasma to about 20×10^6 K The higher temperatures at Princeton were obtained with considerable additional heating provided by injecting into this plasma a stream of energetic neutral deuterium atoms (pre-accelerated as ions to 95 keV and then neutralized)

The additional heating amounted to 17 MW for up to 0 5 s during the latter half of the current pulse, it raised the ion temperature at the centre of the discharge to 230 × 10^6 K This temperature was still rising at the end of the heating period (which is limited by engineering technicalities) The ion number density in the central region (within roughly 20 cm of the centre of the minor cross-section) is of the order of 5 × 10^{19} m^{-3}, and the thermal coupling between the electrons and the ions is relatively slow The electron temperature is about 70 × 10^6 K, noticeably lower than that of the ions The ion temperature was measured by the Doppler broadening of line radiation from nickel and iron impurities (in highly stripped, lithium- and helium-like states), and is consistent with magnetic measurements of the total energy in the plasma (roughly 2 MJ) The energy confinement time, a measure of the thermal insulation obtained by dividing the plasma energy by the input power, is between 0 1 and 0 2 s

The experimental error of the estimated deuterium ion temperature has not been announced, it is probably of the order of 25 per cent, but until a full exposition of the estimation methods is published, some reservation must remain on the value of the temperature quoted The temperature planned for in reactors is about the 230 × 10^6 K (or 20 keV) now claimed No such uncertainty applies to the deuterium–deuterium fusion power in the plasma, which is unambiguously obtained by measuring the neutron flux emitted from the plasma This fusion power is reported to be about 10 kW A substantial but unspecified proportion of this power must be due to direct interaction of the injected 95-keV deuterons with the plasma deuterons, an enhancing effect explicitly sought in the objectives of the Tokamak Fusion Test Reactor If a 50/50 mixture of deuterium and tritium had been used in the same conditions, then the fusion power would have exceeded 1 MW Previously about 0 5 kW of deuterium–deuterium fusion power has been reported from Princeton[3]

The achievement of these results primarily arises from the heating power — the largest yet used on a tokamak As reported both by Princeton and by the Joint European Torus team at a recent symposium[4], neutral injection is an effective way of heating large tokamaks to the 10 keV range, but is nonetheless accompanied by some degradation of the thermal insulation obtained without it, and a consequent two-fold reduction of the energy confinement time, depending on the power used An encouraging feature of the Princeton result is that this degradation has not increased at the higher power used The best heating was obtained using so-called balanced injection that is, the angular momentum imparted to the plasma by the injected deuterium atoms (which, in the unbalanced case, causes the plasma to rotate at near-sonic speeds) is kept small

From the point of view of scientific understanding, the key issue of the magnetic thermal insulation remains The energy losses from the plasma exceed the unavoidable minimum values arising from ion–ion collisions and from the inevitable bremsstrahlung radiation of hot electrons But in some conditions, and in some other experiments, these minimum values have been approached rather closely One encouraging indication given by the Princeton result is that a qualitative characteristic of the ideal thermal insulation, the bootstrap current[5], in which the collisional outward diffusion of the plasma across the magnetic field can by itself sustain the discharge current, has been indicated in some conditions of their reactor

Overall, the preliminary report from Princeton indicates the achievement of conditions close to those hoped for when construction of the $300-million apparatus was approved in 1976, and suggests that fusion power in the megawatt range is achievable in the planned deuterium–tritium experiments The report gives an encouraging foretaste of the results from experiments on all the large tokamak assemblies ☐

1 Zarnstorff M C *et al* *Bull Am Phys Soc* (in the press)
2 Owens D K *et al* *Bull Am Phys Soc* (in the press)
3 Murakami M *et al* *Plasma Phys cont Fusion* **28** 17 (1986)
4 Pease R S Rebut P-H & Bickerton R J *Phil Trans R Soc* (in the press)
5 Bickerton R J *et al* *Nature phys Sci* **229**, 110 (1971)

R S Pease is UKAEA Programme Director for Fusion at the Culham Laboratory Abingdon, Oxfordshire OX14 3DB, UK

Mass extinctions

Catastrophic vegetation changes

Margaret E Collinson

DINOSAURS and ammonites are probably the best-known examples of the many species that became extinct at the Cretaceous–Tertiary (K–T) boundary[1,2] Because there are unusually high concentrations of iridium in the clays at the boundary, it is possible that the biotic changes were caused by asteroid impact[3,4] Terrestrial phenomena such as vulcanism may, however, have been responsible for the extinctions[5–7] Although ecological disruptions of marine biota at this time are widely documented[1–6], little is known about the influence of the K–T event on terrestrial species[8] New evidence from fossil land plants (refs 9, 10 and page 148 of this issue[11]) reveals a widespread ecological trauma, coupled with floral turnover, which occurred at the same time as the marine disruptions

Late Cretaceous floras were dominated by flowering plants in at least four distinctive floral provinces[12,13] In the Northern Hemisphere the Aquillapollenites province includes eastern Siberia and western North America, and the Normapolles province included Europe and eastern North America In the pollen and spore record several features characterize the K–T event There was a major floral turnover in the Aquillapollenites province, with extinction of many forms, more restricted turnover in the Normapolles province, but little or no turnover in Southern Hemisphere provinces A marked increase in the development of modern pollen types occurred in the early Tertiary[8,12,13]

The first detailed evidence of the vegetational expression of these floral changes, the 'fern spike', was found during a study of the K–T boundary in the Raton Basin of the western interior of North America[14] Destruction of pre-boundary angiosperm-dominated vegetation was followed by establishment of a fern-dominated, coal-forming mire vegetation Many Cretaceous flowering plants became extinct, although some persisted across the boundary Re-establishment of angiosperm dominance followed "within a few centimetres of the boundary"[14] but with very different floral composition

These events are known to have occurred throughout the western interior of North America and in Canada[9] Floral change in angiosperm groups falls into three categories major extinctions but with some species persisting across the boundary, groups showing evolutionary changes between pre- and post-boundary forms, and groups largely unaffected at the boundary In the first group, different species become extinct in the northern and southern parts of the region, and macrofloral analysis[11] reveals differential extinction in the southern (up to 75 per cent of Cretaceous species) compared with the northern (up to 25 per cent) part of the western interior The presence of a post-boundary, fern-dominated, coal-forming, mire vegetation was confirmed by the macrofossil analyses

A recent palynological study[10] from a marine sequence in Japan reveals a fern spike at the same time as the major faunal turnover in marine biota This study not only extends the geographical range over which the ecological disruption in the Aquillapollenites province has been recognized, but demonstrates, for the first time, that the marine and terrestrial events happened at the same time

A relationship between leaf form and prevailing conditions has been observed, where vegetation is in equilibrium with climate Using this relationship, Wolfe and Upchurch[11] report in this issue additional aspects of the K–T boundary events in western North America These include an increase in precipitation across the boundary with the initiation of rain forest environments in the early Tertiary, little overall change between late Cretaceous and early Tertiary temperatures, and a high survival rate of deciduous versus evergreen taxa across the boundary both in northern conifer-dominated and in southern angiosperm-dominated vegetation An increase in coniferous elements in the post-boundary vegetation occurred in the Montana/Wyoming area of the western interior of North America[11] and also in Japan[10] (It may also occur in the western Mediterranean if magnetostratigraphic correlation of the sections studied proves correct[15])

Vegetational changes across the boundary have been compared with the destruction and renewal of modern vegetation after catastrophic volcanic eruption[14,16,17] Replacement climax vegetation may, in such cases, be controlled by the disruptive effects of continued disturbance, without reaching equilibrium with the prevailing climate[18] Thus, palaeoclimatic inferences based on K–T boundary vegetation sequences assume that no subsequent disturbance affected post-boundary vegetation succession Sedimentological or other evidence is needed to confirm this assumption

To explain the differential extinction of evergreen taxa, Wolfe and Upchurch[11] suggest a 1–2 month low (near 0 °C) temperature change at or near the boundary The magnitude and duration of such an excursion is clearly constrained by the high survival rate of large aquatic ectotherms[11] But it seems possible that seeds of evergreen taxa survived in the seed bank and re-established when conditions became less hostile

Despite possible problems with such details, it seems clear that the vegetation did change across the K–T boundary in the Aquillapollenites province[9–11] Angiosperm-dominated forests were destroyed and a fern-dominated, coal-forming mire vegetation became established immediately post-boundary at the same time as marine biotic changes Turnover in angiosperm floras varied according to taxon and geographical area, but included both major extinctions and persistence more or less unchanged across the boundary Post-boundary vegetation was therefore different in many ways from that pre-boundary In some areas conifers increased in abundance and selection for deciduousness had a profound effect on Northern Hemisphere but not Southern Hemisphere vegetation[11]

If volcanism in the Indian Deccan Traps did cause the extinctions, as has recently been suggested[19], it is indeed surprising that the Southern Hemisphere provinces show no evidence of the ecological trauma of the Aquillapollenites province Detailed examination of K–T boundary sections from these provinces is clearly needed The last appearance of the dinosaurs may well have preceded these vegetational changes[8], but of the many possible reasons for the extinction of the dinosaurs, those involving a short cold period or enforced changes in herbivore diet bear reinvestigation □

1 Raup, D M & Sepkoski, J J *Science* **231**, 833-836 (1986)
2 Raup, D M *Science* **231**, 1528-1533 (1986)
3 Alvarez, L W, Alvarez, W, Asaro, F & Michel H V *Science* **208**, 1095-1098 (1980)
4 Hsu, K J *et al* *Science* **216**, 249-256 (1982)
5 Officer, C B & Drake, C L *Science* **219**, 1383-1390 (1983)
6 Officer, C B & Drake, C L *Science* **227**, 1161-1167 (1985)
7 Wolbach, W S, Lewis, R S & Anders, E *Science* **230**, 167-170 (1985)
8 Hickey, L J *Nature* **292**, 529-531 (1981)
9 Tschudy, R H & Tschudy, B D *Geology* **14**, 667-670 (1986)
10 Saito, T, Yamanoi, T & Kaicho, K *Nature* **323**, 253 256 (1986)
11 Wolfe, J A & Upchurch, G R *Nature* **324**, 148-152 (1986)
12 Batten, D J in *Fossils & Climate* (ed Brenchley, P) 127-164 (Wiley, Chichester, 1984)
13 Muller, J *Ann Mo bot Gdn* **71**, 419-443 (1985)
14 Tschudy, R H, Pilmore, C L, Orth, C J, Gilmore, J S & Knight, J D *Science* **225**, 1030-1032 (1984)
15 Ashrof, A R & Erben, H K *Palaeontographica* **200**, 111-163 (1986)
16 Leahy, G D, Spoon, M D & Retallack, G J *Nature* **318**, 318 (1985)
17 Spicer, R A, Burnham, R J, Grant, P G & Glicken, H *Fern J* **75**, 1-5 (1985)
18 Cross, A T & Taggart, R E *Ann Mo bot Gdn* **69**, 676-734 (1980)
19 Courtillot, V, Besse, J, Vandamme, D, Jaeger, J -J & Montigny, R *C r hebd Seanc Acad Sci Paris* Ser II **303**, 863-868 (1986)

Margaret E Collinson is in the Department of Biology, King's College London (KQC), Kensington Campus, Campden Hill Road, London W8 7AH, UK

Dating and correlating the French Mousterian

SIR—We note the thermoluminescence dates from Le Moustier[1] with interest and consider them to be valid and informative in the context of that site However, it is regrettable that in his News and Views piece Paul Mellars[2] has used the publication of these dates as an opportunity to reinterpret the chronostratigraphy of Combe Grenal and, by inference, the early Upper Pleistocene sequence of south-west France

Mellars claims that various lines of chronostratigraphic evidence from "key sites" such as Le Moustier and Combe Grenal now converge to suggest that the Mousterian archaeological sequences are successive rather than parallel Unfortunately, this claim is based on the contentious assumption that data from certain so-called "key" sequences can be simply used to interpret other sites, it omits relevant evidence and it balances perilously on the unjustified correlation of the oxygen isotope record from deep sea core V19–30 with arboreal pollen from Combe Grenal The correlations suggested by Mellars fail to take into account a conventional radiocarbon date and thermoluminescence dates for Combe Grenal[3] and ignore significant stratigraphic non-sequences between layers 55–56 and 35–36 at Combe Grenal and H1 and G4 at Le Moustier The proposed correlation between the sites would imply an average error of 50 per cent in the available thermoluminescence dates Although it is clear that the chronostratigraphic interpretation of the Combe Grenal data is not straightforward[3], an error of this magnitude seems most unlikely (S G E Bowman, personal communication) Indeed, the dates tend to support the alternative hypothesis of considerable overlap between the Le Moustier and Combe Grenal sequences as suggested by Bordes[4] and Laville[5,6]

As an alternative to the thermoluminescence dates, Mellars compares arboreal pollen data from Combe Grenal with the oxygen isotope sequence This correlation is problematic in the absence of independent dating to indicate the chronological position and duration of the pollen sequence Without such dating the comparison of curves that represent ice volume and percentage of arboreal pollen rather than actual temperature is speculative, especially as the latter may vary due to local conditions and taphonomic factors as well as climatic change[7,8]

Mellars uses the correlations in his two figures to support his particular view[9] of the succession of archaeological industries from Ferrassie to Quina and then Mousterian of Acheulian Tradition This view would seem to be supported by his positioning of the La Ferrassie hominid from Peyrony's layer D[10,11] at 60,000–75,000 BP This is surprising because it requires a hiatus in the order of 30,000–40,000 years between layers D and E, assuming the generally accepted dating of 30,000–35,000 BP for the Lower Perigordian archaeological assemblage in layer E Further, it does not allow for the possibility that the burials were inserted into layer D and could postdate both the deposits and, in Mellars view, the Ferrassie assemblage Elsewhere[11] the burials have been said to be late Mousterian, dating from Wurm II at around 37,000 BP On the evidence available it would seem unwise to try to arrange the fossil hominid record in line with the model given by Mellars

Overall, the viewpoint presented seems premature and needs to be considered further in the context of a detailed review of the chronostratigraphic and archaeological evidence Otherwise, to paraphrase Bordes[12] on the subject, we remain at risk of setting Charlemagne on a motorcycle

NICK ASHTON
JILL COOK

Department of Prehistoric and Romano-British Antiquities, The British Museum, London WC1B 3DG, UK

1 Valladas, H , Geneste, J M , Joron, J L & Chadelle, J P *Nature* **322,** 452 (1986)
2 Mellars, P *Nature* **322,** 410 (1986)
3 Bowman, S G E & Sieveking, G de G *Pact* **9,** 253 (1983)
4 Bordes, F *A Tale of Two Caves* (Harper & Row, New York, 1972)
5 Laville, H *World Archaeol* **4,** 323 (1973)
6 Laville, H , Rigaud, J-P & Sackett, J R *Rockshelters of the Perigord Geological Stratigraphy and Archaeological Succession* (Academic, New York, 1980)
7 Woillard, G M *Nature* **281,** 558 (1979)
8 Turner, C in *Palynologie Arthéologique* (eds Renault-Miskovsky, J , Bui-Thi, M & Girard, M) 347 (CNRS, Paris 1984)
9 Mellars, P *Proc Prehist Soc* **35,** 134 (1969)
10 Capitan, L & Peyrony, D *Rev Archéol* **22,** 29 (1912)
11 Delporte, H (ed) *Le Grand Abri de la Ferrassie* (Institut de Paléontologie Humaine, Paris, 1984)
12 Bordes, F *Bull Soc Préhist Francaise* **78,** 77 (1981)

MELLARS REPLIES—First, I was well aware of the small series of thermoluminescence dates (6 against the 34 for Le Moustier) produced by the British Museum laboratory for the Combe Grenal sequence[1], but these have little if any direct bearing on the central issues The laboratory itself has already expressed serious reservations over the interpretation of several of these dates[1] and drawn attention to the obvious conflict between the two dates of 105,000 and 113,000 years obtained for the basal levels of the sequence and the totality of the geological, palaeobotanical, faunal and archaeological evidence which points to a *minimum* age for these deposits in the region of 130,000–140,000 BP (= ocean core stage 6, or 'Rissian' in the French terminology)[1–3] Dating of a parallel series of samples by the same laboratory from the nearby site of Pech de l'Azé IV has been dismissed by the laboratory as "too young to be acceptable"[4] For this reason, exact interpretation of the Combe Grenal sequence on the basis of the thermoluminescence dates alone would clearly be hazardous

What is significant is that 5 of the 6 thermoluminescence dates for the Combe Grenal sequence are substantially earlier than any of the 34 dates for the Le Moustier succession — which is fully consistent with the general conclusion that the greater part of the geological and archaeological sequence at Combe Grenal is entirely earlier than that at Le Moustier Note that dating of a sample from the upper (but not final) Mousterian levels at Pech de l'Azé IV yielded a date of 19,600 ± 1,600 BP, which is approximately half of the known age of these levels, and demonstrably in error by a factor of at least ten standard deviations on the measured sample age[4]

Second, I am not quite sure I understand the objections to the proposed correlation between the Combe Grenal sequence and the oxygen isotope record in ocean cores, which is based not on my own interpretations but on the correlations already adopted by Laville and others for the major part of the Combe Grenal succession[5,6] The interpretation of the ocean core sequence as a record of climatic fluctuations depends not only on the record of oxygen isotope ratios but also on studies of faunal assemblages in ocean core sediments (as a reflection of sea-surface temperatures)[7] and on the close correlations now established with palaeobotanical records from terrestrial and offshore sites, particularly the long and well-dated pollen sequence from Grande Pile in north-eastern France[8,9] Similarly, the interpretation of the climatic sequence at Combe Grenal depends not only on changing arboreal pollen ratios (though these are the easiest to illustrate in simple graphical form) but on the parallel and mutually consistent records of geology, sedimentology and faunal assemblages[2,3,10] The remarkable similarity in all these independent records of climatic change can hardly be dismissed by a passing reference to "local and taphonomic factors" and provides unambiguous support for the climatological/chronological correlations proposed by myself, Laville[5], Dennell[6] and others for the Combe Grenal succession

Third, limited depositional hiatuses should perhaps have been indicated in figures 1 and 2 of my article between layers G and H at Le Moustier and layers 35 and 36 at Combe Grenal — but these episodes have no bearing on the specific correlations proposed between the two sites

Fourth, I fully accept the implications of a major hiatus between the Mousterian and Upper Palaeolithic levels at La Ferrassie Similar depositional/erosional hiatuses are well documented at several other cave and rock-shelter sites in south-west France, for example between the

Mousterian and Solutrian levels at Le Placard, and between the Mousterian and Aurignacian levels at Abri Caminade[2,11] Positive indications of this hiatus can be seen in recently published sections of the site[14] The suggestion that the Neanderthal burials at La Ferrassie represent later instrusions into the Mousterian levels is intriguing, but difficult to evaluate, 80 years after the original excavations

Finally, Cook and Ashton make no allowance for the wealth of purely archaeological evidence bearing on the wider problems of Mousterian chronology and intercorrelations between sites Any attempt to support the stratigraphic correlations of Laville[2,12] would need to explain why at least 12 sites in south-west France show a consistent sequence of Mousterian of Acheulian tradition above Quina/Ferrassie industries, why Mousterian of Acheulian tradition industries invariably occur at the top of Mousterian sequences in cave and rock shelter sites, and why Mousterian of Acheulian tradition industries are entirely lacking from all except the uppermost 4 of the 55 levels of Mousterian occupation at Combe Grenal[13] No explanation has so far been offered as to how these observations can be reconciled with the hypothesis of a close synchronism between the archaeological sequences at Le Moustier and Combe Grenal

PAUL MELLARS

Department of Archaeology,
Downing Street,
Cambridge CB2 3DZ, UK

1 Bowman, S G & Sieveking, G de G *Pact* 9, 253 (1983)
2 Laville, H , Rigaud, J-P & Sackett, J R *Rock Shelters of the Perigord Geological Stratigraphy and Archaeological Succession* (Academic, New York, 1980)
3 Bordes, F , Laville, H & Paquereau, M-M *Actes de la Société Linnéene de Bordeaux* 103B, 1-19 (1966)
4 Bowman, S G E *et al Pact* 6, 362 (1982)
5 Laville, H *et al Bull Inst Géol Bassin Aquitaine* 34, 219 (1983)
6 Dennell, R *Nature* 301, 199 (1983)
7 Sancetta, C *et al Quat Res* 3, 110 (1973)
8 Woillard, G & Mook, W G *Science* 215, 159 (1982)
9 Turon, J-L *Nature* 309, 673 (1984)
10 Bordes, F & Prat, F *L'Anthropologie* 69, 31 (1965)
11 de Lumley, H (ed) *La Préhistoire Française* (CNRS, Paris, 1976)
12 Laville, H *World Archaeol* 4, 323 (1973)
13 Mellars, P A *Proc Prehist Soc* 35, 134 (1969)
14 Delporte, H (ed) *Le Grand Abri de la Ferrassie* (Institut de Paléontologie Humaine, Paris 1984)

What can AIDS virus codon usage tell us?

SIR—Macromolecular sequence comparison is extraordinarily useful in elucidating evolutionary relationships, largely because divergence rates are approximately constant, and the chance of convergence is very small The close evolutionary relationship of the human immunodeficiency virus (HIV) and visna virus has been well characterized by this approach[1,2] Grantham, largely responsible for demonstrating that patterns of codon choice are species are species specific[3], and Perrin now suggest that comparison of such patterns in HIV and visna may provide a useful approximate evolutionary distance[4]

Decomposition of a DNA sequence into a codon usage table loses a great deal of information — comparisons of the intact sequences of homologous genes and proteins must be more informative than comparisons of codon choice In fact, the distance measure used by Grantham and Perrin (how many of 18 amino acids have the same most-preferred codon in two codon usage tables) can yield some strange results For example, among the viruses in Table 1, it is thought that Moloney murine leukaemia virus (MMLV, another retrovirus) is the most closely related to HIV, and yet HIV codon usage is most similar to that in the influenza virus, and even cauliflower mosaic virus (CAMV) is more similar to HIV than is MMLV Also, a compilation of 66 mouse genes[5] yields the same preferred codons as the human (total) for each of the 18 amino acids, and yet differs from 60 rat genes[5] for two amino acids Thus, unlike divergence in protein and DNA sequences, differentiation at the level of codon usage is clearly not linear with time It is also highly susceptible to convergence

Table 1 Similarities of codon choice between organisms

	(*N*)	HIV	E LO	E HI	Baci	Y LO	Y HI	H AT	H CG
Human (total)	(135)	2	8	8	1	2	7	3	14
Human G+C rich	(8)	0	8	10	1	0	7	1	
Human A+T rich	(8)	12	8	5	13	14	7		
Yeast HI	(38)	5	2	11	6	11			
Yeast LO	(66)	11	6	6	13				
Bacillus	(21)	11	9	7					
E coli HI	(42)	2	8				5	(3)	MMLV
E coli LO	(58)	6				10	1	(28)	AD2
					3	6	12	(6)	SV40
				11	1	5	13	(8)	Flu
			10	7	5	8	8	(2)	CAMV
			Flu	SV40	AD2	MMLV	HIV		

Values are the number of amino acids (out of 18) where the two data sets show the same preferred codon *N* is the number of genes in each data set Yeast, *Saccharomyces cerevisiae*, *Bacillus*, *B subtilis* HI and LO (for *E coli* and yeast) indicate groups of genes identified as having high and low codon bias Human GC-rich α-actin, apolipoprotein A1, β-chorionic gonadotropin, α-globin, ζ-globin, metallothionein II, insulin, α-tubulin Human AT-rich albumin, α-amylase (pancreatic and salivary), factor IX, γ-fibrinogen, γ-interferon, HPRT, parathyroid hormone Viruses HIV, 5 genes (*gag, pol, sor, env, P27*) of ARV-2, Moloney murine leukaemia virus, Flu, influenza virus isolate B/Lee/40, SV40, simian polyoma, CAMV, cauliflower mosaic virus, AD2, adenovirus 2 All codon usage data was taken from ref 5, except yeast[6], *Bacillus*[10], *E coli*[7] and HIV (calculated from GenBank)

Furthermore, it is clear that there is considerable heterogeneity in codon choice among genes within species In *Escherichia coli* and yeast (the two best studied organisms) groups of genes with high and low codon bias can be identified[3,6,7] that differ somewhat in their preferred codons (Table 1) In mammals there are also large differences among genes in the direction of codon preference[8] These differences seem to be related to local genomic G+C content[8], and can yield very different sets of preferred codons

Thus, when Grantham and Perrin use their comparison to estimate the extent to which viruses imitate the codon preference of their hosts[4], and point out that codon usage in HIV genes is different to that in a compilation of all human (excluding immune system) genes, this is somewhat misleading In fact, HIV genes show considerable similarity in codon choice to a group of human genes with low G+C content Comparing species totals, HIV codon usage shows more similarity to that in *Bacillus subtilis* and yeast, than to that in the compilation of human genes

On the positive side, it is possible that the differences among mammalian genes in pattern of codon usage may partly reflect the diverse tissues in which those genes are expressed One investigation of this possibility[9] gave negative results, but examined very few genes It may be that codon usage in HIV genes reflects adaption to expression in particular tissues, in which case the report[4] that there is some similarity in this respect between HIV genes and T-cell receptor genes would be rather interesting

PAUL M SHARP

Department of Genetics,
Trinity College
Dublin 2, Ireland

1 Sonigo, P *et al Cell* 42, 369-382 (1985)
2 Yasunaga, T , Sagata, N & Ikawa, Y *FEBS Lett* 199, 145-150 (1986)
3 Grantham, R *et al Nucleic Acids Res* 8, r49-r62 (1980), 9, r43-r47 (1981)
4 Grantham, R & Perrin, P *Nature* 319, 727-728 (1986)
5 Maruyama, T , Gojobori, T , Aota, S & Ikemura, T *Nucleic Acids Res* 14, r151-r197 (1986)
6 Sharp, P M , Tuohy, T M F & Mosurski, K R *Nucleic Acids Res* 14, 5125-5143 (1986)
7 Sharp, P M & Li, W-H *Nucleic Acids Res* 14, (in the press)
8 Ikemura, T *Molec Biol Evol* 2, 13-14 (1985)
9 Hastings, K E M & Emerson, C P *J molec Evol* 19, 214-218 (1983)
10 Ogasawara, N *Gene* 40, 145-150 (1985)

Colonization of space: destiny or folly?

Von R Eshleman

The US National Commission on Space has grandiose plans for the colonization of the Solar System. Has the view that unmanned missions may bring much greater benefit to mankind been passed over?

THE presidentially-appointed National Commission on Space has recently released an expansive agenda for the US space programme in the first quarter of the twenty-first century (*Pioneering the Space Frontier*, Bantam Books, Inc., New York, 1986). With space enterprise, exploration and science as components, the overall rationale behind this vision of the future is founded on an assumption that the human species is destined, within the foreseeable future, to colonize space.

But this assumption is not critically examined. It is offered as an act of faith, as though it were rendered self-evident by references to Columbus and 'The New World', the settlement of the American West, and the pioneering spirit that has served us so well in the past.

These are indeed powerful images, but do they necessarily make space colonization a human imperative? I do not question whether it could be done so much as to wonder why it is assumed that it should be done. I would worry about unresolved difficulties, but my main concern is the danger of precluding alternative possibilities of greater promise.

The commission envisages that by the year 2000, the United States would have an Earth Spaceport, with crew, for storage, fueling, repair, supply, maintenance, and launch of transfer and orbital manoeuvering vehicles for access to more distant locales. By the year 2025 there would be three additional spaceports with human crews, including one near Mars, staffed spaceships continually cycling between the orbits of Mars and Earth, manufacturing at a permanent human base on the Moon, astronaut outposts on Phobos and Deimos, the small moons of Mars, and an exploratory base on the surface of Mars which would become a fully-staffed permanent base within a few years. It is stated that this 'Highway to Space' and 'Bridge Between Worlds' would provide the infrastructure for space applications and science, while serving as a focal point for international cooperation and advances in technology. And it would initiate the colonization of the Solar System.

Will this proposal for pioneering the space frontier prove to be prophetic or visionary? It is intended to help influence future events so as to become the former, but current events and their origins suggest that it could well be the latter. It is my opinion that the commission's proposals, if initiated, would lead to a continuation of the fundamental failure of the US space programme to use national resources most effectively for progress at the frontier of space. The crux of the problem, I submit, is an inappropriate mix of the relative influences of our space ends and means and sometimes dreams.

In the report, for example, it is the dream of colonization that determines the proposed infrastructure of spaceports, cycling spaceships and lunar and martian bases. But these are simply means, the ends are primarily the exploration and commercial and scientific use of the space environment. Is there not a danger that the means will become ends in themselves? The question is not even asked the other way around. If all manner of space manufacturing, Earth remote sensing, space communications, planetary exploration, and space science laboratories and observatories are indeed our main ends, by what means may they best be realized? This inevitably brings up the decades-old debate concerning the relative importance and roles of manned and unmanned space endeavours. In the commission's view, the debate is finished. Human crews, settlers, and vacationers belong wherever they conceivably can go. In this scenario, the next frontier will be conquered as before, with a wave of human pioneers.

Ultimate frontier

But mankind could expand into the ultimate frontier of space in another way. Never before have we had the ability to observe, explore, utilize, examine, monitor, manipulate, prospect, mine and manufacture, or in short pioneer and operate at a new physical frontier, without the necessity of a physical human presence at that frontier. In a way, humanity as a whole could experience its continuous expansion into thousands of locales at the frontier of space, with a multitude of its doers (like the hand of Viking 1 as it clutched a sample from an ancient plain of Mars) and sensors (like the eyes of Jason Jr as it descended the grand staircase of the Titanic or the magnetometers of Voyager 2 as they charted the bizarre field of Uranus) on the ends of thousands to billions of kilometre-long radio nerve fibres and commandable sinews. Remote human monitoring and control of space manufacturing and delicate experimentation could also be accomplished.

The means to secure these ends would be low-cost unmanned launch and transfer vehicles and new levels of capabilities in robotics, communications and control, adaptive and expert systems, and many related technologies. Our collective space dreams could encompass the endless possibilities for international cooperation in obtaining new knowledge, products and resources from space in this way, and the endless potential of applying them to better the human condition on Earth. Such dreams include an automation revolution to rival the industrial revolution in significance, and new levels of understanding of climatic effects like carbon-dioxide warming, nuclear winter and acid rain from remotely controlled examinations of related phenomena on other worlds. Concerns about national prestige and demonstrations of technological prowess, as with Apollo, or dreams about solving international tensions by a joint symbolic venture in space, such as a human mission to Mars, do not appear to me to be useful guides for progress.

The current hiatus in American access to space due to the Challenger disaster should be used to re-appraise our goals and the methods for their realization. It is clear that the present problems would have surfaced in some way, even if the tragic accident had not occurred. Although many observers have cheered the decision by President Reagan to replace Challenger, some seem to have missed the significance of the accompanying announcements that the United States will phase out commercial and foreign payloads for the shuttle fleet, and will encourage the commercialization of launch services after the 15-year delay that was enforced by nationally-subsidized competition. This stands as a tacit admission that we have failed to realize the fundamental rationale for the space shuttle programme, namely, that of providing low-cost access to space for all US and many foreign users.

The shuttle experience stands as one example of the kinds of things that could go wrong if we try to bypass incremental, well-funded, long-term progress based on carefully defined goals and feasible, cost-effective means tailored to those ends. I submit that space colonization is a dream whose time has not yet come, and that the benefits of the space frontier beckon us to realize them in the most direct and effective ways possible. □

Von R Eshleman is Director of the Center for Radar Astronomy, DB221, Stanford University, Stanford, California 94305, USA

The quark plasma

H. Satz

Fakultat fur Physik, Universitat Bielefeld, Postfach 8640, D-4800 Bielefeld 1, FRG
and Physics Department, Brookhaven National Laboratory, Upton, Long Island, New York 11973, USA

Because nucleons are bound states of quarks, nuclear matter should at sufficiently high density become a quark plasma. Assuming quantum chromodynamics as the basic theory of nuclear forces, it is possible to derive strong-interaction thermodynamics and study the nature of this transition. Predictions of thermodynamic variables are obtained by computer simulation; they provide the first theoretical information on the phase structure of strongly interacting matter. The experimental search for the quark plasma is beginning this year at Brookhaven and at CERN. Energetic nuclear collisions could produce sufficiently dense and long-lived strongly interacting systems to allow a first look at the predicted new state of matter.

IN the past two decades our picture of an elementary particle has changed significantly Today we consider protons, neutrons and all the other 'elementary' particles in the realm of nuclear forces as composite as bound states of quarks Elementary particles are no longer the smallest units of matter, they, in turn, consist of quarks The binding force between quarks is quite different, however, from the forces usually encountered in physics: at short distances it becomes so weak that the quarks appear to be free, but with increasing distance of separation, it grows without limit Thus it is not possible to split an elementary particle into its quark constituents—this would require an infinitely large force Elementary particles remain the smallest units of matter that can exist by themselves, individually, quarks can only exist together with other quarks or antiquarks

This modified view of the elementary particle concept has in turn had a profound effect on our picture of nuclear matter at high density If nucleons (protons and neutrons) were elementary hard spheres, then a densely packed system of such objects would be the limiting state of matter (see Fig 1) For composite nucleons, however, interpenetration will occur if the density is increased enough, and eventually each quark will find very many others in its immediate neighbourhood Because it has no way to remember which of these were its partners in the low-density nucleonic state, the concept of a nucleon and of nuclear matter has become meaningless here Thus it is expected that at a high enough density a new state of matter will appear the quark plasma—a system whose basic constituents are unbound quarks

Before considering in more detail how quark binding can be dissolved in dense matter, let us get some idea of the densities involved Dense atomic solids contain about 10^{23} nucleons cm^{-3} nuclear matter, as it exists in the interior of neutron stars, has about 10^{38} nucleons cm^{-3}, and is thus 10^{15} times denser than matter in the world around us What is the effect of further compression in this regime of extreme density?

I have already noted that the strength of the binding force between quarks increases indefinitely with their separation, making the 'ionization' of a nucleon into its quark constituents impossible the quarks are 'confined' Why then should they become deconfined in sufficiently dense matter? The answer to this question is found in charge screening, which already occurs in atomic solids The force between the charged partners of a bound state is considerably modified, if this bound state is placed in an environment of many other such objects The Coulomb potential between two electric charges e, separated by a distance r, has in the vacuum the form e^2/r, in the presence of many other charges it becomes subject to Debye screening, $e^2/r \rightarrow (e^2/r)\exp(-r/r_D)$, where the screening radius r_D is inversely proportional to the overall charge density of the system the more other charges present in the environment, the shorter the range of the force between two given charges If in atomic matter the Debye radius r_D becomes less than the atomic radius r_A, then the binding force between electron and nucleus is effectively screened, and the electron becomes 'free' For atomic systems, an increase in density thus results in an insulator-conductor transition[1,2] the binding between charges is dissolved in the presence of enough other charges

Screening, in contrast to ionization, is thus a short-range mechanism to dissolve bound states, and for this reason we expect it to apply also to quark binding I shall return shortly to the theoretical basis of the force between quarks, note here only that the interaction is again based on a charge, the 'colour' charge, which in strong-interaction physics plays the role of the electromagnetic charge in atomic physics Nucleons are colour-neutral bound states of colour-charged quarks, and hence nuclear matter corresponds to a colour insulator With increasing density, colour screening should set in and lead to the transition from colour insulator to colour conductor to the quark plasma The predicted new state of matter thus is a macroscopic system which is conductive to the basic charge of the strong interaction

The theoretical analysis of strongly interacting matter is based on the fundamental theory for the interaction of quarks, quantum chromodynamics I shall survey the present state of this analysis in the next section The experimental investigation of these phenomena is the main aim of forthcoming studies of

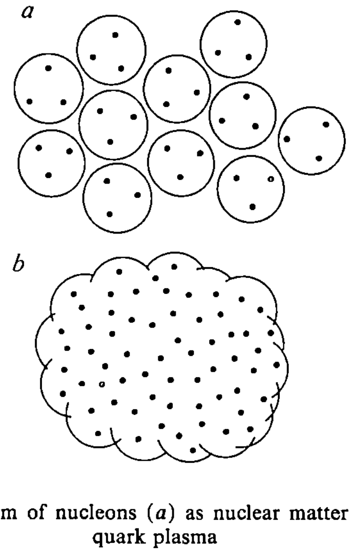


Fig 1 A system of nucleons (*a*) as nuclear matter and (*b*) as quark plasma

ultrarelativistic collisions of heavy nuclei, this will be taken up in the section after next

Statistical quantum chromodynamics

Quantum chromodynamics (QCD) describes the interaction of quarks in much the same way as quantum electrodynamics (QED) does for electrons Spinorial matter fields ψ—here quarks, electrons in QED—interact through massless vector gauge fields A, here gluons take the place of photons Just as in QED, the interaction in QCD is determined by an intrinsic charge, the 'colour', each quark can be in one of three different colour charge states In contrast to QED, however, the gauge fields here also carry a charge each gluon can be in one of eight colour states As a result, the gluons, as well as the quarks, interact with each other, and it is in fact this gauge-field interaction that is the crucial feature of QCD Gluonic lines of force between two quarks now attract each other, and this turns the three-dimensional Laplace equation, with the Coulomb solution $V(r) \sim 1/r$, into a one-dimensional form, with $V(r) \sim |r|$ as 'confining' solution

The Lagrangian density of QCD has the form[3-5]

$$\mathscr{L} = -\tfrac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a - \sum_f \bar{\psi}^a_f (\imath - g)_{\alpha\beta} \psi^\beta_f \qquad (1)$$

with

$$F^a_{\mu\nu} = (\partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^a_{bc} A^b_\mu A^c_\nu) \qquad (2)$$

Here $a, b, c = 1, \quad , 8$, and $\alpha, \beta = 1, 2, 3$ denote the gluon and quark colour charges, respectively, they are based on SU(3) as the colour gauge group, whose structure constants are f^a_{bc} The explicit summation in equation (1) runs over the possible different species ('flavours') f of quarks, I restrict myself here to the (practically massless) u and d quarks, which suffice to construct all non-strange nucleons and mesons

Note again that it is the last term in equation (2), and hence the non-abelian nature of the gauge group[6], which constitutes the crucial feature of the theory, for $f^a_{bc} = 0$, equations (1) and (2) have the same structure as QED

The Lagrangian (1)/(2) contains a single, dimensionless coupling g As a result, QCD cannot predict any dimensional observable, but only the dimensionless ratios of observables In many ways, such a description comes close to an ideal physical theory it is closed in itself, and contains no outside parameters As shown below, QCD thus predicts for instance the transition temperature from nuclear matter to quark plasma in units of the nucleon mass, and this mass is itself calculated within the theory as the lowest bound-state energy of three massless quarks

Given the Lagrangian density (1)/(2), we have in principle all the dynamical input necessary for strong interaction thermodynamics From $\mathscr{L}$, we must determine the Hamiltonian H, and then construct in the usual way the partition function

$$Z(T, V) = \mathrm{Tr}(\exp(-H/T)\} \qquad (3)$$

for a thermodynamic system at temperature T and within a spatial volume V, the trace in equation (3) runs over all microscopic physical states From Z, we obtain by differentiation the desired thermodynamic observables, getting for instance $\varepsilon = (T^2/V)(\partial \ln Z/\partial T)_V$ for the energy density, and $P = (\partial \ln Z/\partial V)_T$ for the pressure But how is it possible to carry out the actual evaluation of Z? The perturbative approach familiar from QED could be applied here as well for the regime in which the interactions are weak—that is, at very high density But the central theme of statistical QCD, the transition from the strongly coupled nuclear matter to the weakly coupled plasma, is by nature non-perturbative and cannot be addressed in such a fashion

The only approach known at present that permits us to study strongly interacting matter at all densities is the lattice formulation introduced by Wilson[7] Replacing the space-time continuum, on which the fields ψ and A are defined, with a discrete and finite set of lattice points, one obtains a form of the partition function that is divergence-free and, moreover, structurally very similar to that of spin systems in statistical mechanics It is therefore possible to use the computer simulation techniques developed for such spin systems in the evaluation of lattice QCD[8] Renormalization group theory, which studies the scale-dependence of statistical systems, then permits us to recover the continuum limit (vanishing lattice spacing) of the results thus obtained In the past few years this method of evaluation has been applied extensively and with considerable success in the study of the phase structure of strongly interacting matter (for a recent survey, see ref 9)

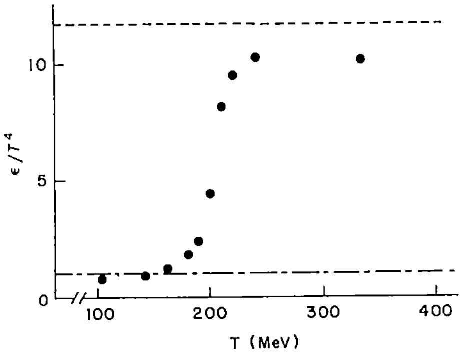


Fig. 2 The energy density of strongly interacting matter as a function of the temperature compared with the energy density of the ideal quark-gluon plasma (---) and of the ideal pion gas (- - -) (From ref 12)

After taking a brief look at the computer simulation of statistical systems[10], I shall here go on directly to the results obtained so far

The analytical mathematics available today has been developed to a large extent for isolated systems, even very simple many-body systems with interaction cannot generally be solved, once the interaction really connects the whole system and is not perturbatively reducible to small parts of it As an example, recall that the three-dimensional Ising model is not analytically solved In recent years, with the advent of large-scale computers, it has become possible to simulate such systems the given system is reproduced on the computer in a specific microscopic configuration In reproducing all possible configurations, it would then be possible to evaluate $Z(T, V)$ exactly for systems of finite size Because the three-dimensional Ising model—N^3 spins pointing either up or down—already leads to 2^{N^3} configurations, it is evident that for meaningful systems such a procedure is not feasible, even on the largest computers Therefore, in statistical physics Monte Carlo methods have been developed to generate 'typical' configurations[11], and, averaging over a sufficiently large number of these for the simulated system, it is possible to 'measure' the thermodynamic quantities of interest

The density of strongly interacting matter can be increased in two distinct ways by compression, or by heating, which owing to the relativistic energy range involved leads to particle production and hence a higher density So far most of the results of statistical QCD deal with the latter of these two cases, with 'hot mesonic matter' of vanishing baryon number density The other extreme, the compression of 'cold nuclear matter', is now the subject of much study, the computer simulation of this system is technically more difficult

Beginning with matter of vanishing baryon number density, the energy density of a non-interacting gas of massless quarks and gluons is

$$\varepsilon_P = (37\pi^2/30) T^4 \simeq 12 T^4 \qquad (4)$$

Just as in the Stefan-Boltzmann form for the photon gas, the numerical factor in equation (4) is determined by the number of degrees of freedom of the constituents involved their spins, colours and flavours Similarly, a gas of non-interacting π mesons becomes, if we neglect the meson mass, $\varepsilon_\pi = (3\pi^2/30)T^4 \simeq T^4$, taking into account the three possible charge states of the π In the low-temperature region, where quarks are confined, we might expect values of ε in this range, whereas at high enough temperatures the plasma form (4) should become correct The energy density obtained by computer simulation of statistical QCD[12] is shown in Fig 2, there is a very rapid transition from the mesonic regime to the plasma The transition temperature—converted to physical units through comparison with the mass of the nucleon or the ρ meson—is around 200 MeV, which in turn means that an energy density of at least 2 5 GeV fm^{-3} is necessary to create a quark-gluon plasma

Before we turn to nuclear matter, a further aspect related to the deconfinement of quarks should be noted An electron in a conductor has an 'effective' mass different from that in a vacuum or in an atom This mass shift is the consequence of collective effects of all the positive ions and other electrons present Similarly, we can consider an 'effective' quark mass, which in the confined phase is around 300 MeV (corresponding to the three constituent quarks of a nucleon), in the weakly interacting plasma it vanishes This shift in the effective quark mass need not exactly coincide with deconfinement In principle it may occur later, which would lead us to a three-phase structure for strongly interacting matter hadronic matter, consisting of mesons and nucleons, a constituent quark plasma, consisting of massive but deconfined quarks, and finally the quark-gluon plasma of massless quarks and gluons For systems of vanishing baryon number density, however, statistical QCD has shown that the two phenomena indeed do occur at the same point[12,13]

Systems of non-vanishing baryon number density—'nuclear' or 'baryonic' matter—are now a subject of considerable attention The partition function depends in this case on a further variable the baryonic chemical potential μ, which through $n_B = (T/V)(\partial \ln Z(T, V, \mu)/\partial\mu)$ determines the overall baryon number density of the system (the number of baryons minus that of antibaryons, per unit volume) So far, numerical studies have been made only for rather small values of μ[14,15] The resulting phase diagram is shown in Fig 3, up to this point, deconfinement and quark mass shift still coincide Whether that will remain true even at very low temperatures, or whether the mentioned additional phase, the constituent quark plasma, makes its appearance there, remains to be seen

Nuclear collisions at high energies

The thermodynamics of strongly interacting matter is of importance to cosmology Very soon after the Big Bang, the Universe must have been a hot, dense quark-gluon plasma, its rapid expansion presumably reduced the density such that after about 10^{-3} s the transition to the hadronic Universe occurred Only from that point on was there a physical vacuum

In recent years the possibility of recreating this cosmological process in the laboratory has come within reach[16] The collision of two heavy nuclei at high enough energies should produce a 'little bang', providing a bubble of strongly interacting matter sufficiently large and of sufficiently long lifetime to allow the application of statistical QCD, that is, it should form the quark plasma, and make possible the study of the transition to hadronic matter What is the physical basis for this?

When two nucleons collide at high energy, as for example in one of the large proton accelerators at CERN (Geneva) or Fermilab (Illinois), they almost always just pass through each other In the process the nucleons are excited, but in addition they leave a 'condensation trail' of deposited energy in the region of space they passed through (see Fig 4) This energy is rapidly transformed into hadrons, the secondary particles usually produced in such collisions From the analysis of proton-proton collision data, it is known that the deposition of energy at high enough energies is about 1/3 GeV fm^{-3} It does not increase much more with a further increase in collision energy, which only makes the condensation trail longer

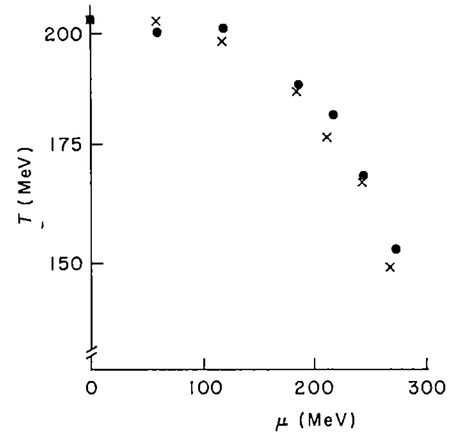


Fig 3 The dependence of the deconfinement temperature (×) and the chiral-symmetry restoration temperature (●) on the baryonic chemical potential μ, increasing μ means increasing the baryon number density (From ref 15)

If instead of the nucleons, two heavy nuclei collide (for simplicity, both with mass number A), a multiple superposition of the phenomenon just described is obtained, so that a given region of the condensation trail now receives a much larger deposition of energy[17] Simple geometrical arguments lead to an energy density of at least $\varepsilon \sim A^{1/3}$ GeV fm^{-3}, this means that an energetic collision of two ^{238}U nuclei provides an average deposition of about 6 GeV fm^{-3} As seen in the previous section, this is well above what is needed for plasma formation

Processes of this type can also be simulated on large computers, using simple phenomenological interaction models as input, One such stimulation (done by D Strottman (Los Alamos), based on a model he developed with A S Goldhaber (Stony Brook)) is shown in Fig 5, we see clearly the formation of the condensation trail after the nucleons have penetrated each other The process shown here is the collision of a ^{16}O nucleus with a^{238}U target, at an incident energy of 100 GeV per nucleon Such reactions will be the first to be studied in forthcoming experiments at CERN (Geneva), I shall return to the experimental situation shortly

Nuclear collisions should thus produce bubbles of rather high energy density, and of a volume $V_A \sim A$ fm^3, although not really macroscopic, V_A is, for large A at least, much bigger than the typical hadron volume of about 1 or 2 fm^3 It is therefore hoped that such a bubble represents a system to which the equation of state obtained in statistical QCD may be applied, if the energy

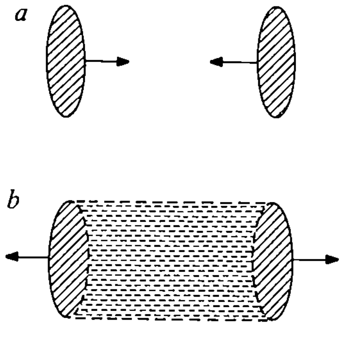


Fig 4 High-energy collision of two nucleons (a) just before, (b) just after the collision

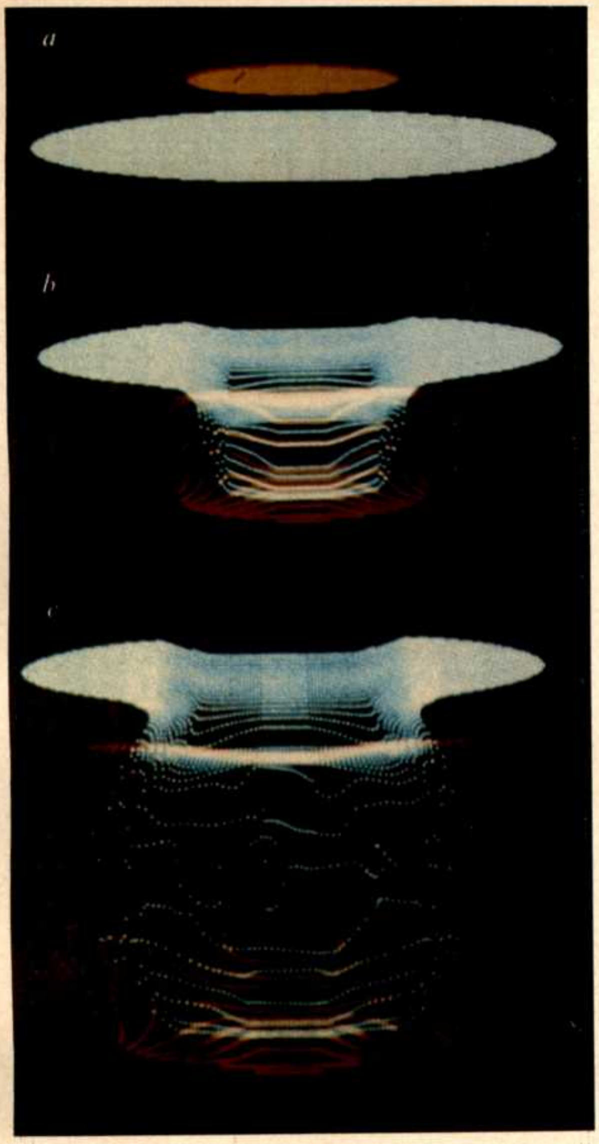


Fig. 5 Successive stages of a high-energy nuclear collision: ^{16}O at 100 GeV per nucleon on ^{238}U; (*a*) at approximately 10^{-24} s before the collision, (*b*) at 2×10^{-23} s and (*c*) at 10^{-22} s after the collision. (From the computer simulation by D. Strottman.)

deposit is high enough, the initial temperature will be high, and the bubble will contain quark matter.

Each plasma bubble will now expand quite rapidly and in doing so cool down to the deconfinement transition temperature. Hadrons—predominently mesons—will form, and after still further expansion, at the energy density of about 1/3 GeV fm^{-3}, they will cease to interact and form the secondary particles produced in the collision. The plasma expansion has been a subject of much recent work[18]; it is usual to apply the hydrodynamic flow equations, using an equation of state of strongly interacting matter (ideally that obtained from statistical QCD) as input.

Given this model for plasma formation and transition in nuclear collisions, how can it be tested experimentally? What are the signatures for the creation of a new state of matter in such interactions? What has emerged so far is a number of different indications, which, if all observed together, would make plasma formation the most likely explanation.

At the time when any measurement is made, the 'little bang'—provided it occurred—is over. Hence one thing to look for is a signal from the initial plasma phase that is not affected by the subsequent transition. If quarks in the plasma, by collision or annihilation, emit photons, then these, because of the much weaker electromagnetic interaction, will generally escape from the system without further interactions. The spectra of photons, real or virtual (as e^+e^- pairs), should thus tell us something about the early stage of the produced system.

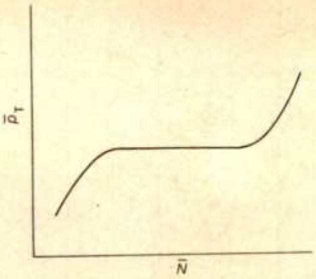


Fig. 6 Schematic dependence of the average transverse momentum $\bar{p}_T$ of hadronic secondary particles on the average number $\bar{N}$ of such secondary particles, as expected for energetic nuclear collisions leading to a first-order deconfinement-confinement transition. (From ref. 19.)

Because the number of produced secondary hadrons is directly related to the initial energy density, the multiplicity in a given collision may serve as an indicator for ε. In the condensation trail we find bubbles of various momenta in the longitudinal direction (that is, along the collision axis). Hence the longitudinal momenta of the secondary particles are not directly related to any specific plasma bubble. The average transverse momenta $\bar{p}_T$ can be so related, however, and thus provide a measure of the initial temperature: the higher p_T, the greater the momentum boost due to the plasma expansion must have been, and hence the higher the initial temperature *T*. Thus, using $\bar{p}_T$ as a measure of temperature and the multiplicity $\bar{N}$ in a given momentum range as a measure of energy density, from Fig. 2, the $\bar{N}$ dependence of $\bar{p}_T$ should be as shown in Fig. 6[19]. Indeed, there are first indications for such a behaviour, obtained from the study of nuclear collisions in data from cosmic rays[20].

For a survey of experimentally observable signatures see ref. 21. As this article was being completed, a further and apparently rather clear signal for plasma formation was developed[22]. If deconfinement occurs, then the production of J/Ψ mesons in nuclear collisions must be very strongly suppressed, because within the plasma the charmed quark constituents of the J/Ψ cannot form a bound state, and the large mass of the J/Ψ prevents it from being formed in a later stage. No other mechanism for such a suppression is yet known. It should be emphasized, however, that with the study of dense hadronic systems we enter an unexplored region of physics, and thus surprises are to be expected.

The experimental study of high-energy nuclear collisions is scheduled to start this year. At Brookhaven National Laboratory a beam of ^{32}S ions at 15 GeV per nucleon should become available in October 1986, and three major project groups plan to perform experiments with this beam. At CERN a beam of ^{16}O (perhaps also ^{32}S) ions, of energies up to 200 GeV per nucleon, will be used by five major experimental groups in a three-week run starting in November 1986. In both laboratories there will be additional smaller-scale emulsion experiments. It is certainly hoped that these studies may provide first indications of the quark plasma; nevertheless in both cases the ions are not very heavy, and *A* should be large to obtain as 'macroscopic' a system as possible. At Brookhaven a booster facility is under construction, which will allow the acceleration of ions as dense as gold (^{197}Au); it should go into operation in around 1989. Beyond these projects, which rely on existing accelerators both at BNL and CERN, there are plans for the construction of a machine dedicated to high-density hadron physics at Brookhaven. This relativistic heavy ion collider (RHIC) would provide a colliding nuclear beam facility, with energies of around 100 GeV per nucleon in each beam for even the heaviest ions. A first decision on this proposal is expected soon.

Why do a rapidly growing number of physicists find these prospects so exciting? It appears to be within reach to study the states of matter in strong interacting physics, and in this study to recreate experimentally in the laboratory a crucial stage in the development of the early Universe.

This work was supported by the US Department of Energy contract DE-ACO2-76CH00016.

1. Mott, N. F. *Proc. phys. Soc.* **A62,** 416–422 (1949).
2. Mott, N. F. *Rev. mod. Phys.* **40,** 677–683 (1968).
3. Fritzsch, H., Gell-Mann, M. & Leutwyler, H. *Phys. Lett.* **47B,** 365–368 (1973).
4. Weinberg, S. *Phys. Rev. Lett.* **31,** 494–497 (1973).
5. Gross, D. J. & Wilczek, F. *Phys. Rev.* **D8,** 3633–3652 (1973).
6. Yang, C. N. & Mills, R. L. *Phys. Rev.* **96,** 191–195 (1954).
7. Wilson, K. *Phys. Rev.* **D10,** 2445–2459 (1974).
8. Creutz, M. *Phys. Rev.* **D21,** 2308–2315 (1980).
9. Satz, H. *A. Rev. nucl. Part. Sci.* **35,** 245–270 (1985).
10. K. Binder (ed.) *Monte Carlo Methods in Statistical Physics* (Springer, Berlin, 1979).
11. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. *J. chem. Phys.* **21,** 1087–1092 (1953).
12. Çelik, T., Engels, J. & Satz, H. *Nucl. Phys.* **B256,** 670–686 (1985).
13. Kogut, J., Polonyi, J., Wyld, H. W., Shigemitsu, J. & Sinclair, D. K. *Phys. Rev. Lett.* **53,** 644–647 (1984).
14. Engels, J. & Satz, H. *Phys. Lett.* **159B,** 151–154 (1985).
15. Berg, B., Engels, J., Kehl, E., Waltl, B. & Satz, H. *Z. Phys.* **C31,** 167–174 (1986).
16. Kajantie, K. (ed.) *Quark Matter 1984, Lecture Notes in Physics* no. 221 (Springer, Berlin, 1985).
17. Bjorken, J. D. *Phys. Rev.* **D27,** 140–151 (1983).
18. Baym, G. *Quark Matter 1984, Lecture Notes in Physics* no. 221 (ed. Kajantie, K.) 39–52 (Springer, Berlin, 1985).
19. Van Hove, L. *Phys. Lett.* **118B,** 138–140 (1982).
20. Miyamura, O. *Quark Matter 1984, Lecture Notes in Physics* no. 221 (ed. Kajantie, K.) 187–195 (Springer, Berlin, 1985).
21. Gyulassy, M. *Nucl. Phys.* **A418,** 59c–85c (1984).
22. Matsui, T. & Satz, H. *Phys. Lett.* **178,** 416–422 (1986).

ARTICLES

Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*

Hans Georg Frohnhöfer & Christiane Nüsslein-Volhard

Max Planck Institut für Entwicklungsbiologie, Spemannstrasse 35, 74 Tübingen, FRG

Embryos derived from bicoid (bcd) *females lack head and thorax. Cytoplasmic transplantation experiments reveal that* bcd^+*-dependent activity is localized at the anterior pole of wild-type embryos. The activity can induce anterior development in mutant embryos at any position along the antero-posterior axis and suppresses posterior development.*

WHEN the young *Drosophila* embryo proceeds through the stages of syncytial blastoderm towards gastrulation, the perivitelline layer of somatic nuclei looks almost homogeneous, with no morphological signs of cellular specification. But when the blastodermal layer is examined with molecular probes, the spatially restricted transcription of segmentation genes reveals that it is already being subdivided into segmental units[1-3]. It has frequently been proposed that maternally encoded functions provide positional information in the egg and control the spatial pattern of zygotic gene expression, which in turn determines the developmental fate of individual cells. In insects, there is evidence for two centres of activity, located at the egg poles, which organize the antero-posterior axis (reviewed in ref. 4). In *Euscelis* transposition of posterior polar plasm within the egg can induce a duplicated posterior end with inversed polarity[4]. An analogous activity was found at the anterior in chironomid midges; manipulation of the anterior pole can cause double abdomen formation (reviewed in ref. 5). The activity appears to be required for anterior and to inhibit posterior development. The anterior activity in *Smittia* is sensitive to RNase and ultraviolet irradiation[5] and are therefore characterized as RNA-containing particles.

Evidence for terminal organizing centres has been obtained in *Drosophila* by experimental removal and/or transplantation of polar cytoplasm[6]. Specific, distinct phenotypes are produced which resemble the phenotypes of several maternal effect mutants. This suggests that the genes involved are important in the establishment of the morphogenetic centres. Five loci are known in which mutations eliminate the posterior pole plasm, including the polar granules. No pole cells are formed and the development of the abdomen is abolished. Strikingly, the posteriormost somatic region of the embryo, the telson, is not affected[7-9]. In *oskar* mutants (osk^-), abdominal segmentation can be restored by transplantation of wild type pole plasm into the prospective abdominal region[9]. Only pole plasm from osk^+ embryos can induce posterior development at the anterior, suggesting that this group of genes is involved in the organization of posterior pattern[6,9].

The anterior structures, head and thorax, are eliminated by maternal effect mutations of the bicaudal group[10-12]. Head and thorax are replaced by a second duplicated abdomen (including telson) of reversed polarity. Bicaudal embryos are produced by mutations in a variety of complementation groups with quite different genetic properties, but in no case with 100% penetrance[12]. Therefore these mutants do not readily define functions specifically required for anterior development. Here we describe the maternal effect gene *bicoid* (*bcd*). Embryos produced by mutant females lack head and thorax and have a second duplicated telson at their anterior end, but usually no abdominal segments of inversed polarity. We show that the function of *bcd* is mediated by a cytoplasmic activity, localized at the anterior, with long range organizing effects on the antero-posterior pattern.

Phenotype of *bicoid*

All eleven *bcd* alleles are strictly maternal. The penetrance is 100%, but the alleles differ in the extent of anterior pattern defects they produce (Fig. 1). Embryos from strong mutant alleles completely lack the head and thorax (Fig. 1*d*); instead they have a second telson at the anterior end. Deletions and fusions of segments are also found in the anterior abdominal region. Occasionally abdominal segments with inversed polarity are duplicated at the anterior, together with the telson. The weakest alleles show pattern defects only in the anterior head

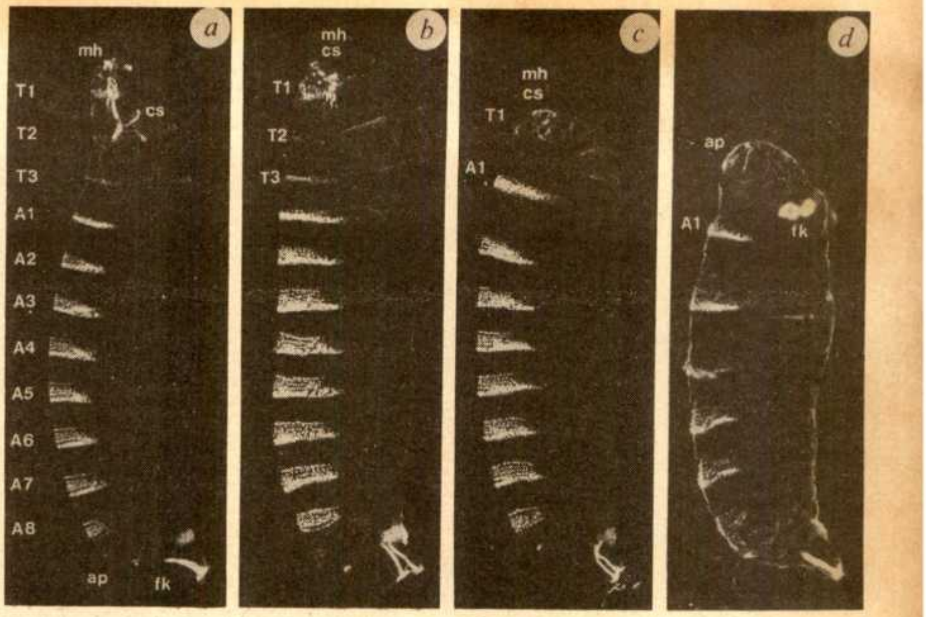


Fig. 1 The phenotype of *bicoid*. Cuticular preparations of embryos of different mutant strength. *a*, Wild type. The weakest *bcd* alleles ($bcd^{2\text{-}13}$, bcd^{E5}, and bcd^{111}) lack only labral derivatives (median tooth, dorsal bridge) and are not shown. *b*, Weak intermediate phenotype (bcd^{E3}/bcd^{E3}, 18°). The entire head is reduced; thoracic and abdominal segmentation is normal. *c*, Strong intermediate phenotype (bcd^{E3}/bcd^{E3}, 25°). The head is further reduced and the thorax is affected. On the ventral side denticle belts T2 and T3 are deleted and T1 is reduced in size. Three dorsal bulges with segmental arrays of dorsal hairs demonstrate that the thoracic region is still composed of all three segments fused together. Intermediate alleles: bcd^{E3} (strongly temperature-sensitive), bcd^{E4} and bcd^{085}. *d*, Strong *bcd* phenotype (bcd^{E1}/bcd^{E1}). Head and thorax are lacking altogether. They are replaced by a complete telson with anal plates, tuft, spiracles and filzkörper, but these are not stretched. The anterior abdomen shows segment defects and fusions. Strong alleles: bcd^{E1}, bcd^{E2}, $bcd^{33\text{-}5}$, $bcd^{23\text{-}16}$, bcd^{GB}. A more detailed description of the individual alleles and their temperature dependence will be given elsewhere. Mutant alleles: *Bicoid* alleles were induced with ethyl-methanesulfonate in a marked chromosome and isolated: (1) in a large scale mutagenesis experiment designed for recessive female sterile mutations on the third chromosome (C.N.-V., G. Jürgens, K. Anderson, R. Lehmann & H.G.F., unpublished, see also refs 9, 24) (bcd^{E1}-bcd^{E5}, bcd^{085}); (2) as alleles which failed to complement bcd^{E3} ($bcd^{2\text{-}13}$, bcd^{111}, $bcd^{33\text{-}5}$, bcd^{GB}) or (3) from individual F_1 females heterozygous for a mutagenized chromosome, on the basis of the dominant haploinsufficient phenotype (anterior shift of the head fold) ($bcd^{23\text{-}16}$, *Df*(3*R*)*LIN*). The *bcd* gene is not complemented by *Df*(3*R*)*Scr*, *Df*(3*R*)*LIN*, or *Df*(3*R*)9*A*99, but covered by *Df*(3*R*)Ns^{R+17} and *Df*(3*R*)4*SCB* which places it in 84A into the same cytogenetic interval which contains *zerknüllt*[25-27].

Methods. Embryos were collected on yeasted agar plates, dechorionated and devitellinized. They were mounted in Hoyers mounting medium, diluted 1:1 with lactic acid and cleared by overnight incubation at 60 °C. Darkfield photographs. A1-A8 = ventral abdominal denticle bands, ap = analplates, cs = cephalopharyngeal skeleton, fk = filzkörper, mh = mouth hook, T1-T3 = thoracic denticle bands. Anterior at the top, ventral at left. For the morphology of the larval cuticle see refs 28, 29.

region. With increasing allelic strength the entire head region (Fig. 1*b*), then the thorax (Fig. 1*c*) are affected. Otherwise there is no clear order in which structures are eliminated. Three anterior regions with different sensitivity for bcd^+ reduction can be distinguished; the preantennal head region, antennal and gnathocephalic head segments, and thorax.

In transheterozygous combinations, *bcd* deficiencies behave like strong alleles. The allelic series thus reflects a continuous decrease of bcd^+ activity (hypomorphic states); the strong alleles approximate the complete loss of bcd^+ function (amorphic condition).

Development

Bicoid embryos develop normally until blastoderm formation. Morphological alterations in mutant embryos start at gastrulation. The cephalic fold, indicating the position of the gnathal anlagen, is a good landmark for changes of organization (Fig. 2). Females carrying a duplication or a deletion for bcd^+ produce embryos with the cephalic fold shifted towards a more posterior or anterior position respectively (Fig. 2*a*–*c*). With further decrease of bcd^+ activity in homozygous alleles, the cephalic fold becomes weaker, and finally disappears (Fig. 2*d*). At intermediate mutant strength a posterior midgut plate is formed at the anterior end, indicating a duplicated telson. This is seen more clearly in strong alleles (Fig. 2*e*). However, no pole cells are formed at the anterior end and no pole plasm is found there in the early stages. The shifted cephalic fold reveals that the *bcd* phenotype is associated with changes in the blastodermal anlagen plan. Thus the reduction in the size of anterior structures in the differentiated embryo can be traced back to shrinkage of the corresponding primordia. The abdominal anlage is enlarged and spreads towards the anterior (data not shown).

Localization of bcd^+ activity

We tried to restore the *bcd* function in strong mutant embryos by transplantation of cytoplasm from wild type embryos. Cytoplasm was taken from the anterior tip of young (stage 2)[13] wild type embryos and transplanted into the anterior of mutant embryos of the same age. The amount of transplanted material corresponded to about 5% of the egg volume; one wild type embryo was used as a donor for 1–2 mutant recipient embryos. Nearly all recipient embryos showed a change in phenotype from strong *bcd* towards normal. Some embryos formed only thoracic segments (Fig. 3*b*); others, in addition to thorax, gained almost complete heads (Fig. 3*a*). Three exceptional embryos survived until pupation. The anterior abdomen, which is misshapen in the mutant, became normalized whenever thorax was induced. However, the normal number of abdominal segments was often not restored; new defects arose in the midabdominal region. The embryos were classified according to the degree of anterior development observed (Fig. 4).

The rescuing activity of anterior cytoplasm is strongly dependent on the genotype of the female producing the donor eggs (Fig. 4*b*). Cytoplasm from bcd^- embryos has no rescuing activity. Eggs laid by females with only one bcd^+ gene (*Df* bcd^-/bcd^+), the normal two doses (bcd^+/bcd^+) and three doses of the gene (*Dp* $bcd^+/bcd^+/bcd^+$) have cytoplasm with increasing bcd^+ activity. Varying the amount of transferred cytoplasm mimicks the effect seen with varying bcd^+ gene dosage (Fig. 4*a*), suggesting a linear relationship between bcd^+ gene expression and transplantable activity. This suggests that the bcd^+ gene product might be the active transplanted factor and that no negative feedback acts on its accumulation at the anterior end.

The efficiency of phenotypic rescue decreases with increasing age of both donors and recipients (Fig. 4*c*, *d*). Unfertilized eggs

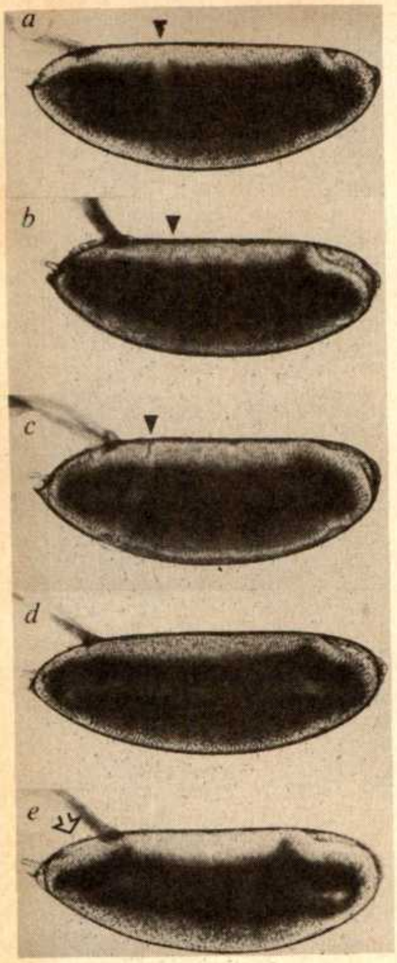


Fig. 2 Gastrulation pattern of mutant embryos. Photographs of living embryos produced by females with: *a*, three bcd^+ copies ($C(1)RM.y$; $T(Y;3)MA9^P/T(Y;3)\ A109^D$)[30,31]; *b*, two bcd^+ copies (wild type); *c*, one bcd^+ copy ($bcd^{GB/+}$); *d*, bcd^{E4}/bcd^{E4}, intermediate allele; *e*, bcd^{E1}/bcd^{E1}, a strong allele. The position of the cephalic furrow (arrowhead) varies with bcd^+ activity between 62% EL (*a*), 65% (*b*) and 71% (*c*). The significance of the backshift in *a* is obscure because females carrying a duplication tend to produce eggs with an unusual shape. Both types of embryos with shifted cephalic folds usually produce normal patterns and survive to adulthood. The cephalic fold is absent in intermediate and strong alleles (*d*, *e*). In the strong phenotype (*e*), a second posterior midgut plate is evident at the antero-dorsal side of the embryo (open arrow). Later in development it performs characteristic invagination movements. Embryos were collected on agar plates, covered with Voltalef 3S oil and photographed in transmitted light with a Zeiss photomicroscope. Anterior left, dorsal at the top.

contain the most active cytoplasm. The activity in older embryos decreases slowly until the last cleavage, after which the activity is sharply reduced (Fig. 4*c*). The high activity observed in unfertilized eggs persists for much longer (data not shown). The responsiveness of recipients decreases continuously (Fig. 4*d*), suggesting that the bcd^+ activity is involved in a process which lasts from fertilization to the end of cleavage.

To determine the localization of the *bcd*-dependent activity in wild-type embryos, cytoplasm was taken from various positions each covering about 10% egg length (EL) along the antero-posterior axis. Figure 4*e* shows that bcd^+ activity is high at the anterior pole and declines sharply away from it, suggesting a graded distribution of bcd^+ activity in the anterior third of the egg. However, the resolution of 10% EL is at the limit of our technique, and a high homogeneous distribution from 85–100% EL would also fit the data. No dorso-ventral difference in activity and no clear difference between central yolk and cortical cytoplasm was obvious. In older donors (syncytial blastoderm), the prevalent anterior localisation is still present (data not shown).

Induction of anterior development

Anterior structures can be induced by injecting bcd^+ activity at any position along the antero-posterior axis. The anterior structures are formed near the injection site. The corresponding reorganization of the pattern involves major changes in embryonic polarity. The abdomen also reorganizes according to the conditions installed by the anterior determinants. The ability of the injected material to induce new thoracic or cephalic structures strongly decreases with distance from the anterior pole (Fig. 5*f*). The posterior pole seems to exert a suppressive force which declines toward the anterior end.

We transferred increased amounts of cytoplasm to compensate for reduced efficiency in induction in the posterior half of the egg. The results are shown together with those of standard experiments in Fig. 5*a*–*f*. We now describe some typical examples of the phenotypes created by heterotopic transplantation, starting with the phenotypic response upon anterior injection.

Even embryos rescued by anterior injection display slight deviations from the phenotypes of the allelic series. They tend to have a rather poorly developed head compared to the thorax (Fig. 3*b*). It may be that when the material is transplanted it is not deposited in the very anterior tip but is spread over an area around 85% EL. If so, the distribution of the activity influences the relative sizes of head and thorax. Injections at 67% EL produced about the same range of phenotypes as anterior injection.

Injection at 50% EL induces head structures near the middle of the injected embryo, flanked by two thoracic regions with opposite polarity (Fig. 3*c*, *e*, *g*). Telson structures characteristic of the mutant phenotype are still present at both poles. A maximum of anterior positional information is thus established near the site of injection in the centre of the egg and decreases in both directions toward the poles. Abdominal segments are only formed at the original posterior site and are frequently reduced in number. The phenotypes suggest that the blastodermal region giving rise to abdomen has been compressed into a very posterior region of the egg.

The asymmetry observed in the embryos is also seen in the inverted region of the induced thorax, which tends to have fewer segments than the one with normal polarity. In the borderline case there is no inverted région in the induced thorax, suggesting a general preference to establish anterior maxima at the terminal region of the egg rather than in the middle. Further, the suppression by posterior factors may primarily affect the posterior portion of the injected material and cause anterior maxima to be established at blastodermal positions somewhat anterior to the site of injection.

Injections to more posterior positions (15% EL and 32% EL) can induce head structures together with three thoracic segments at the posterior end of the egg, the polarity of the entire embryo being reversed. At the same time formation of abdominal segments is strongly suppressed, never leaving more than one field of abdominal denticles. Complete suppression of the abdomen is observed more frequently after injection at 32% EL than at 15% EL (Fig. 5*a*–*f*), suggesting that *bcd*$^+$ activity probably does not act on the posterior pole plasm to suppress abdomen specific factors[9]. Rather it appears to act on a signal spreading from the pole or to share common targets with this signal.

When abdominal segments are formed in phenotypically less extreme animals, the posterior segments have inversed polarity. Figure 3*d*, *f* shows an embryo in which, in addition to producing thoracic structures, two inversed abdominal segments were formed, the posterior one having the appearance of a first abdominal segment. As the anterior activity cannot induce abdominal development on its own, it must have modified the action of some abdomen specific factors. Such an interaction would ensure a continuous change of morphological values at the boundary of thorax and abdomen.

Phenocopies by cytoplasm removal

The *bcd* phenotype can also be produced (phenocopied) by elimination of anterior cytoplasm. Pricking the anterior end of wild type eggs (stage 2) to allow leakage of ~5% volume mainly generates embryos with reduced head structures, resembling weak *bcd* phenotypes (Fig. 6*a*). Treating embryos in which *bcd*$^+$ activity is already reduced causes stronger *bcd* phenocopies approaching the amorphic phenotype (Fig. 6*b*, *c*). Strong mutant embryos do not show any change on loss of anterior cytoplasm (Fig. 6*d*). The artificial phenotypes also show abdominal defects suggesting that the entire *bcd* phenotype is caused by the lack of anteriorly localized cytoplasmic activity. Pricking preblastodermal eggs at non-terminal sites does not eliminate cuticular pattern elements, whereas leakage at the posterior pole with removal of the posterior pole plasm causes abdominal defects resembling weak *oskar* phenotypes[6]. In *Smittia*, in contrast to *Drosophila*, anterior pricking creates bicaudal embryos with a reversed abdomen at the anterior[14]. This difference reflects a general tendency in *Smittia* to produce symmetrical states, if one terminal activity is not balanced by its antagonist.

Conclusions

Mutant *bcd* embryos fail to produce head and thorax. Corresponding primordia do not appear in the anlagen plan at the blastoderm stage. Instead a second telson is found at the anterior end and the abdominal primordium is expanded towards the anterior. The *bcd* phenotype is thus composed of two types of defect; loss of anterior structures due to shrinkage of the anlagen

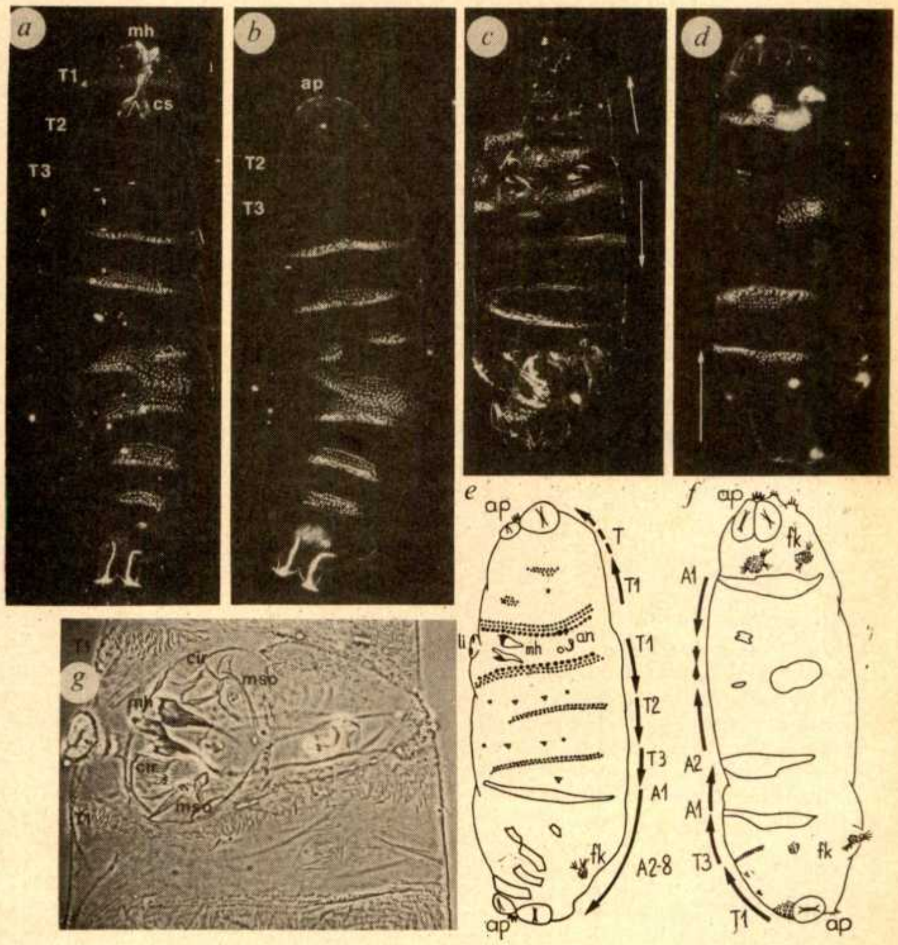


Fig. 3 Phenotypes of *bcd*$^-$ embryos injected with anterior cytoplasm. *a*, *b*, Strong and weak response to anterior injection. The embryo in *a* has gained all anterior structures with the exception of the labral derivatives (median tooth and dorsal bridge). Note that the anterior abdominal region is also normalized; but mid-abdominal segments are fused, a type of defect not seen in non-injected controls. The embryo in *b* developed two thoracic segments, but no head structures; small anal plates are retained at the anterior. In the phenotypic series, embryos with more than one thoracic segment also have some remnants of head structures (compare Fig. 1*c*). Moreover injected embryos display T2 and T3, but the denticle belt of T1 is the most stable to mutation. These discrepancies are interpreted as being caused by a subterminal and less compact positioning of the activity in the injected animals. *c*, *e*, *g*, Embryo after injection at 50% EL. The induced cephalic structures are flanked at the anterior and at the posterior side by thoracic regions of opposite polarity. Most of the head structures are formed at the dorsal side of the embryo (shown in detail in *g*). Mainly maxillary derivatives can be recognized. In this type of embryo, the head is not oriented along the antero-posterior axis of the egg; cephalic structures frequently have dorsal or lateral positions and their sequence seems to be delineated tangentially along the circumference of the egg within a transverse plane. The transverse plane is the plane of bilateral symmetry seen among the head structures. Sometimes pairs of head structures occur at either side of the embryo, (4 mouth hooks, 4 maxillary sense organs and/or 4 groups of ventral cirri in total). Due to cuticular holes and the scarcity of characteristic structures other than those of the maxilla the anteroposterior polarity of the head is difficult to assess. Possibly the anterior end of the head lies either dorsally or ventrally depending on the localization of the injected cytoplasm. *d*, *f*, Embryo injected at 15% EL. Two bands of thoracic denticles have been induced at the posterior. They are the anteriormost elements of an inversed region including about 2.5 abdominal segments. Anteriorly the original abdominal polarity is preserved. The appearance of a new A1 shows that not only the polarity but also the quality of abdominal segments has been changed. The posterior telson is partially suppressed. Arrows, segment polarity; polarity is ambiguous in the case of the broken arrow in *e*. Black dots and triangles, thoracic sensory organs. Identification of labial (li) and antennal (an) structures is not certain in the specimen in *e*. cir, cirri; mso, maxillary sense organ. Other designations as for Fig. 1.

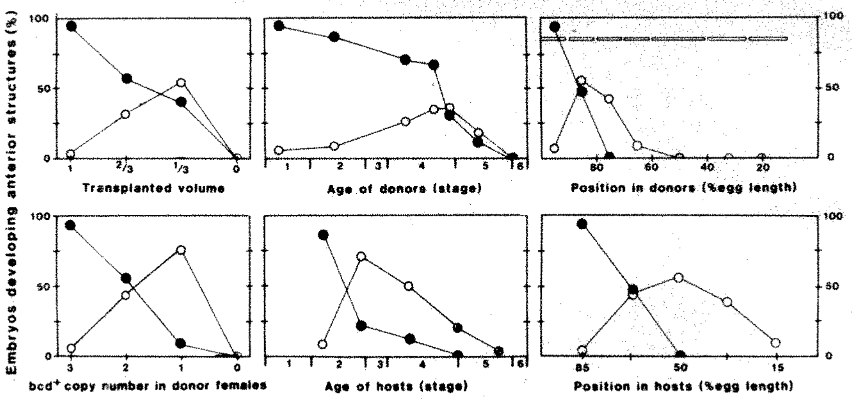


Fig. 4 Phenotypic response on cytoplasmic injection. Frequencies of embryos induced to develop thoracic and cephalic structures (●, strong response) or displaying thoracic structures only (○, weak response) provide a measure for efficiency of induction obtained under various experimental conditions. Females carrying a duplication for bcd^+ produced the donor eggs and the recipients were from bcd^{E1}/bcd^{E2} females unless otherwise indicated. Donor eggs were of stage 1 or 2 in *d* and unfertilized elsewhere; recipient eggs were of stage 1 or 2. Cytoplasm was withdrawn from an anterior area (80–100% EL). When injected into an anterior position it was observed to spread in the central core of the egg between 75 and 95% egg length. The amount of cytoplasm transferred was about 5% of the egg volume. Average number of embryos for each point was 33 (range 13–54). *a*, The usual amount of injected cytoplasm is set to 1 arbitrary unit. The different relative volumes (1, 2/3, 1/3) were achieved by counting the time cytoplasm was allowed to flow from the needle into the egg (15 s, 10 s, 5 s respectively). *b*, Donor females were of the genotypes $bcd^+/bcd^+/bcd^+$ $(T(Y{:}3)MA9^P/T(Y{:}3)\ A109^D)$, wild type, $bcd^-/+(Df(3R)9A99/+)$, and bcd^{E1}/bcd^{E2} with bcd^+ copy number decreasing from three to zero. *c*, *d*, Stages as described in methods. Unfertilized eggs are taken to represent the youngest stage 1 in *c*. In old recipients (stage 5) despite the low frequency of induction affected animals mostly display some remnants of head structure surrounded by few thoracic denticles (●). *e*, Bars on top represent the areas from where cytoplasm was withdrawn. *f*, Further information about the phenotypes obtained in this experiment is provided in Fig. 5*a*–*c*.

Methods. Eggs were dechorionated with NaOCl and attached to a sticky cover slip. Recipient embryos were dried for 10–12 min with a fan, both donors and recipient eggs were then covered with halocarbon oil (Voltalef 10S). Injections were performed under a compound microscope. Micropipettes of about 12 μm external diameter driven by a syringe were used. Recipient eggs were entered by the micropipette at the anterior pole in most cases, sometimes from the side but never from posterior in order to avoid loss of the posterior pole plasm. To assess the amount of cytoplasm transferred, the entire content of an egg was drawn into the needle. About 20 recipient embryos could be injected using one donor embryo. Cytoplasm which leaked out during injection was removed afterwards. Experiments were performed at 18 °C or 20 °C. Usually about 80% of injected embryos differentiated cuticle and could be scored. The stages are those of Campos-Ortega and Hartenstein[13] (1, freshly laid egg; 2, early cleavage; 3, pole cell formation; 4, syncytial blastoderm (nuclear cycle 10–13)[32]; 5, blastoderm cellularization; 6, gastrulation). All embryos were staged visually during injection; young and old stage 2 embryos in *d* were distinguished by the time they entered gastrulation.

and abdominal segmentation defects in a stretched posterior domain. The action of bcd^+ on both the anterior and abdominal region is mediated by the anteriorly localized activity. The anterior centre thus influences the pattern of more than half of the egg length in spite of its striking anterior localization. The structures of the telson occupy the terminal positions in the egg if there is low bcd^+ activity, irrespective of which other structures become their neighbours.

The amount of bcd^+ activity present in the egg determines the size of cephalic and thoracic anlagen. It also affects the degree of anteriorness in the cuticular structures. The exact positioning of the bcd^+ dependent activity seems to influence the relative preponderance of thoracic and cephalic elements in homotopic transplantations. Changes of polarity associated with heterotopic transplantations provide more striking evidence that the bcd^+ activity not only establishes the anterior primordium but also controls its internal organization. Describing the action of *bcd* therefore requires more than one spatial parameter. The production of different anterior qualities depending on quantity and distribution of bcd^+ activity is best explained by assuming a graded distribution of factors determining the positions of head and thorax. The well organized polarity observed in the cuticle of injected embryos contrasts with the coarse conglomeration of activity present immediately after transplantation. The

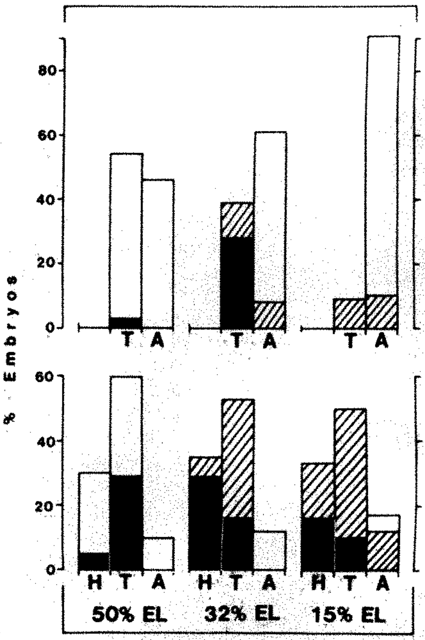


Fig. 5 The effect of heterotopic transplantation on abdomen formation. The standard volume (*a*, *b*, *c*) or 1.5× standard volume (*d*, *e*, *f*) of anterior cytoplasm was transplanted into three different positions at 50% EL (*a*, *d*). 32% EL (*b*, *e*) and 15% EL (*c*, *f*) of bcd^{E1}/bcd^{E2} embryos; other conditions were as described for Fig. 5. On an average 35 embryos have been recorded for each panel. Columns in each panel show the frequency of embryos with head and thoracic structures (H), only thorax (T), or neither of the two (A, abdomen). The marked areas in each column indicate the behaviour of the abdomen occurring in each class. Black portion, embryos with complete abdominal suppression (without abdominal denticle belts); hatching, cases in which inversed abdominal segments have been induced; white, the remainder of the embryos with an abdomen (partially suppressed or not). The phenotype expressed may depend on whether the transferred cytoplasm happens to be closely centred around the injection site or spreads some way through the egg. A certain amount of cytoplasm may be sufficient to induce head if centred well at 50% EL, but will then not suppress the abdomen very strongly. If the same amount of cytoplasm induced only thorax, abdominal suppression was much stronger suggesting less restricted localization (*d*). In many of the embryos displaying thorax only, segmental organization was poor and polarity was disordered. Otherwise we observe the expected positive correlation between transplanted volume, extent of anterior inductions, and suppression of abdomen. Abdominal suppression is most effective at 32% EL, whereas abdominal inversion is obtained with higher frequencies at the posteriormost injection site. The differences between the three sites tend to decrease with higher amounts of transplanted cytoplasm, probably due to lower spatial resolution. Once the threshold of posterior inhibition is overcome, anterior induction may be favoured in terminal positions. On transferring large volumes of cytoplasm, frequencies of polarity reversal within induced anterior regions varied with the site of injection thus: 2% (85% EL), 12% (67% EL), 55% (50% EL), 6% (32% EL), 5% (15% EL).

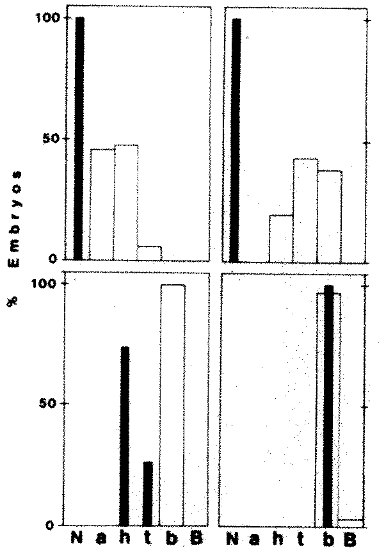


Fig. 6 *a–d*, Production of *bcd* phenocopies by removal of cytoplasm. The phenotypic response upon pricking and subsequent leakage of anterior cytoplasm from different genotypes is shown. Maternal genotype of pricked eggs: *a*, wild type; *b*, $Df(3R)9A99/+$: *c*, bcd^{E3}/bcd^{E3}; *d*, bcd^{E1}/bcd^{E2}. The letters along the abscissa designate the phenotypic classification. N, essentially normal; a, head weakly reduced with antenna still present; h, head strongly reduced; t, thorax reduced (head reduced or absent, usually some telson structures duplicated); b, head and thorax entirely absent, mimicking the bcd^- phenotype; B, some abdominal segments inversed reminiscent to *Bicaudal* (only 1 case). Black columns, phenotypic spectrum of untreated embryos; white columns pricked embryos. Removal of anterior cytoplasm causes a shift toward the amorphic phenotype. Treating amorphic mutant embryos elicits no further change (for example, toward bicaudal embryos).

Methods. Embryos were mounted as in transplantation experiments, but not dried. Anterior pricking caused loss of cytoplasm (5% egg volume), which was carefully removed from the outside of the eggs. About 70% of the embryos survived to the differentiated stage[6]. At least 30 embryos have been recorded for each panel.

transplanted activity may achieve some redistribution, but we feel that it probably acts as a localized source providing a spreading signal to instruct the peripheral sheet of nuclei. This might explain why the position of the transplanted material is crucial for the eventual phenotype but is not necessarily isomorphic with the pattern induced. An anteriorly localized RNA transcribed from the *bcd* locus and distributed similarly to our transplantable activity has been found[15]. The translation product may be the postulated signal.

The spatially restricted expression of zygotic segmentation genes of the 'gap' class might be the focus of *bcd* function in the anterior half of the embryo[3,16]. Gap mutants eliminate continuous blocks of segments (or parasegments[17]) along the body axis[18–20]. In contrast to the *bcd* phenotype, the anlagen plan outside the regions deleted suffers no major alterations[20–22]. In *hunchback* mutants, labium and thorax (the posterior part of the *bcd* domain) are lacking[20]. Therefore there may be no expression of *hunchback* in *bcd* embryos. Mutants in *Krüppel* (*Kr*) eliminate segments T1–A5, segments which are dependent on the function of *bcd*, as well as segments dependent on *oskar*. *Kr* should therefore be expressed in both *bcd* and *osk* embryos. However, the abdominal expansion seen in *bcd* embryos suggests an anterior shift in the spatial expression of *Kr*.

Abdominal development is strictly dependent on maternal genetic functions as exemplified by the gene *oskar*, which has been examined by means quite similar to our own work on *bcd*[9]. In controlling the posterior pole plasm and abdominal development, *oskar* and its relatives[7,8] constitute a posterior analogue to the function of *bcd*. However the interplay between *bcd* and *oskar* dependent factors is not symmetric. Anterior anlagen show no major expansion towards the posterior in *oskar* embryos[9,33], so the establishment of the anterior *bcd* domain is independent of the posterior centre. But an antagonistic effect can be provoked experimentally; when posterior pole plasm is transplanted to the anterior tip of wild-type eggs it suppresses anterior development[6,9]. Anterior cytoplasm transplanted into the range of the future abdomen is weakened in its inductive strength (Fig. 4*f*). Because the anterior source material seems to be outside the range of inhibiting posterior factors in wild-type eggs, there is no obvious natural function for such an interaction during postfertilization stages, when both polar centres are already established.

The bcd^+ gene is not required to establish abdominal anlagen as such but is required to confine abdominal development to the posterior half of the embryo. Depending on position, bcd^+ activity can drive abdominal primordia backwards in the blastodermal anlagen plan; change abdominal segmental qualities toward more anterior values; or suppress abdominal development. Thus bcd^+ exerts profound interventions on structures which develop fairly autonomously in bcd^- embryos. If one assumes abdominal development to be organized by a posterior gradient emanating from the pole plasm[23], then inhibition and modification of this gradient by *bcd* dependent activity would explain all types of interactions described here. Alternatively, *bcd* might not act on maternally encoded factors and contribute nothing to abdominal development other than to determine its anterior boundary by control of *Kr* gene expression.

We thank Ruth Lehmann for discussions and Kavita Arora, Thomas Berleth, Stephen Cohen, Martin Klingler, David Marcey and Siegfried Roth for helpful suggestions on the manuscript. Ruth Lehmann, Kathryn Anderson and Gerd Jürgens participated in the mutant screens that led to the isolation of the *bicoid* alleles. We thank Gunilla Weinraub for drawings, Roswitha Groemke-Lutz for the photographic prints, and Dominique Kaiser for typing the manuscript.

Received 13 August; accepted 8 October 1986.

1. Hafen, E., Kuroiwa, A. & Gehring, W. J. *Cell* **37**, 833–841 (1984).
2. Ingham, P. W., Howard, K. R. & Ish-Horovicz, U. *Nature* **318**, 439–445 (1985).
3. Knipple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A. & Jäckle, H. *Nature* **317**, 40–44 (1985).
4. Sander, K. *Adv. Insect Physiol.* **12**, 125–238 (1976).
5. Kalthoff, K. in *Determinantns of Spatial Organization* (eds Subtelney, S. & Konigsberg, I. R.) 97–126 (Academic, New York, 1979).
6. Frohnhöfer, H. G., Lehmann, R. & Nüsslein-Volhard, C. *J. Embryol. exp. Morph.* (in the press).
7. Boswell, R. E. & Mahowald, A. P. *Cell* **43**, 97–104 (1985).
8. Schüpbach, T. & Wieschaus, E. *Wilhelm Roux Arch. Devl Biol.* **195**, 302–317 (1986).
9. Lehmann, R. & Nüsslein-Volhard, C. *Cell* (in the press).
10. Bull, A. L. *J. exp. Zool.* **161**, 221–241 (1966).
11. Nüsslein-Volhard, C. *Wilhelm Roux Arch. Devl Biol.* **183**, 249–268 (1977).
12. Mohler, D. & Wieschaus, E. *Genetics* **112**, 803–822 (1986).
13. Campos-Ortega, J. A. & Hartenstein, V. *The embryonic development of* Drosophila melanogaster (Springer, Heidelberg, 1985).
14. Schmidt, O., Zissler, D., Sander, K. & Kalthoff, K. *Devl Biol.* **46**, 216–221 (1975).
15. Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. & Noll, M. *Cell* (in the press).
16. Meinhardt, H. *J. Cell Sci.* Suppl. **4**, 357–387 (1986).
17. Martinez-Arias, A. & Lawrence, P. A. *Nature* **313**, 639–642 (1985).
18. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795–801 (1980).
19. Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Devl Biol.* **104**, 172–186 (1984).
20. Lehmann, R. & Nüsslein-Volhard, C. *Devl Biol.* (in the press).
21. Carroll, S. B. & Scott, M. P. *Cell* **45**, 113–126 (1986).
22. Ingham, P. W., Ish-Horowicz, D. & Howard, K. R. *EMBO J.* **5**, 1659–1665 (1986).
23. Meinhardt, H. *J. Cell Sci.* **23**, 117–139 (1977).
24. Anderson, K. V. & Nüsslein-Volhard, C. *Nature* **311**, 223–227 (1984).
25. Lindsley, D. & Zimm, G. *Dros. Inf. Serv.* **62**, 1–226 (1985).
26. Wakimoto, B. T., Turner, F. R. & Kaufman, T. C. *Devl Biol.* **102**, 147–172 (1984).
27. Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Wilhelm Roux Arch. Devl Biol.* **193**, 283–295 (1984).
28. Lohs-Schardin, M., Cremer, C. & Nüsslein-Volhard, C. *Devl. Biol.* **73**, 239–255 (1979).
29. Jürgens, G., Lehmann, R., Schardin, M. & Nüsslein-Volhard, C. *Wilhelm Roux's Arch. Devl Biol.* **195**, 359–377 (1986).
30. Lehmann, R. dissertation, Univ. of Tübingen (1985).
31. Lindsley, D. & Sandler, L. *Genetics* **71**, 157–184 (1972).
32. Zalokar, M. & Erk, I. *J. microsc. Biol. Cell* **25**, 97–106 (1976).
33. Carroll, S. B., Winslow, G. M., Schüpbach, T. & Scott, M. P. *Nature* **322**, 278–280 (1986).

A parallax effect due to gravitational micro-lensing

B. Grieger, R. Kayser & S. Refsdal

Hamburger Sternwarte, Gojenbergsweg 112,
D-2050 Hamburg 80, FRG

The astrophysical importance of gravitational micro-lensing due to stars in the deflecting galaxy is now well-known[1–7]. Each macro-image (caused by the corresponding smoothed-out galaxy) can be split up into several micro-images with typical angular distances measured in micro-arcseconds. Here we discuss a parallax effect due to gravitational micro-lensing (star disturbances). We show that such an effect would most favourably be observed during high amplification events when a compact source (quasar) crosses a critical curve. During such an event even a relatively small displacement (≥0.1 AU) of the observer may lead to measurable differences in the image brightness. We discuss this parallax effect with special emphasis on the determination of (1) the size of the quasar and (2) the relative transversal velocities involved. It should be possible to observe such a parallax-effect with a relatively small telescope (10–20 cm) included in an interplanetary mission. In extreme cases, annual oscillations in the light curves of lensed quasars can be observed from the Earth, but in most cases these oscillations can be neglected.

Of particular interest among instances of gravitational micro-lensing are the high amplification events[2] that occur when a compact source crosses a critical curve[1]. These high amplification events where $\Delta m \approx 3$ are suddenly initiated (or terminated) by the appearance (or disappearance) of two bright micro-images[1,4,5]. It is a kind of 'gravitational eclipse' effect, allowing the formation of large spatial magnitude gradients and the occurrence of very fast changes in the observed brightness due to transverse motion (grad $m \approx 0.3$ AU^{-1} and $\partial m/\partial t \approx 0.1$ day^{-1} for a source radius of 10^{-4} pc and a typical transverse velocity of 500 km s^{-1}). The frequency of such high amplification events has been estimated[2] to be about 0.3 per year for all known gravitationally-lensed quasars. The time scale of such events is given approximately by the diameter of the source divided by the transverse velocity V_T of the source relative to the critical curve in the source plane. We note that V_T is equivalent to a relative transverse velocity (v_T) between the observer and the critical curve in the observer plane

$$v_T = -V_T \frac{D_d}{D_{dq}}(1+z_d) \tag{1}$$

where D_d and D_{dq} are apparent size distances and z_d is the redshift of the deflector. Because the critical curves have a radius of curvature of the order of one normalized length unit[2] ($\sim 10^4$ AU for solar mass stars), we can approximate the critical curves by straight lines inside the solar system. We can also assume that the observed luminosity of a point source locally only depends on the distance d to the critical curve ($\propto d^{-1/2}$)[1,4,8].

During a high amplification event, the observed brightness changes very rapidly, and even a small displacement of the observer can bring about measurable changes. During a typical event ($\Delta m \approx 3$), a displacement of only 0.1 AU (perpendicular to the critical curve) in the plane of the observer leads to a change in brightness of up to $\delta m \sim 0.025$ for a typical lensed quasar ($D_d \approx D_{dq}$ and a source radius of 10^{-4} pc). (For the somewhat untypical lens object OSO 2237+0305 ($D_{dq}/D_d \sim 10$), $\delta m \sim 0.25$.)

This parallax effect raises the possibility of calculating v_T from observations made by three observers some distance apart (see Fig. 1). If the time lags δt_{AB} and δt_{AC} in the light curves are known, we can eliminate the angle β and determine the relative transverse velocity v_T (perpendicular to the critical curve) in the observer plane:

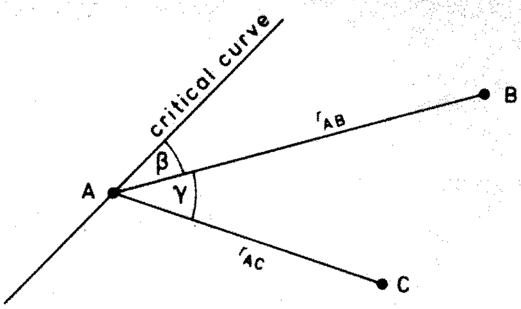


Fig. 1 Position of 3 observers, projected onto the observer plane.

$$v_T = |\sin\gamma| \left\{ \left(\frac{\delta t_{AB}}{r_{AB}}\right)^2 + \left(\frac{\delta t_{AC}}{r_{AC}}\right)^2 - 2\cos\gamma \frac{\delta t_{AB}\,\delta t_{AC}}{r_{AB}r_{AC}} \right\}^{-1/2} \tag{2}$$

Here we have made use of the fact that the time lags δt are equal to the time differences across the critical curve.

From equation (1) we obtain the corresponding velocity V_T in the source plane. From the typical time scale Δt of the high amplification event, we can then determine the time needed by the source to cross the critical curve, and thereby the radius $R = k\Delta t/V_T$ of the source, where k (≈ 1) depends on the luminosity profile of the source and the amplitude of the event. This will be discussed in a subsequent paper.

If there are only two observers, A and B, a straightforward relation between the two unknowns v_T and $\sin\beta$ emerges (see Fig. 1).

$$v_T = r_{AB}\sin\beta/\delta t_{AB} \leq r_{AB}/\delta t_{AB} \tag{3}$$

We thus see that with only two observers an upper limit of v_T is obtained.

Because reliable determination of the time lags does not require very exact observations we propose that a rather small telescope (diameter 10 to 20 cm) would be sufficient in the case of the brightest multiple quasars where $m \approx 17$. In view of the very important information that could be obtained the possibility of sending such a telescope in an interplanetary mission should be investigated.

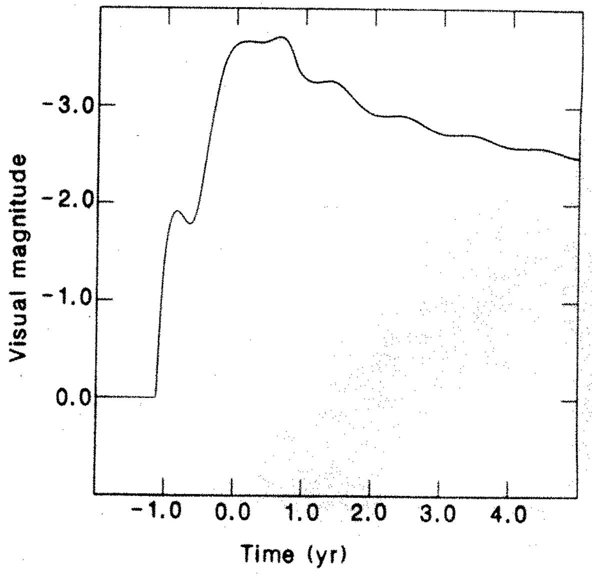


Fig. 2 Simulated light curve of a high amplification event in the lens object 2237+0305, showing oscillations due to the annual motion of the earth.

Considering the sighting of the somewhat improbable lensing by a double star[9,10] in our galaxy ($D_{dq} \approx 10^7\ D_d$), a parallax effect could possibly be observed by Earth-bound observers. During the most rapid variations, an observer displacement of only 6 km results in a brightness change of up to about 0.1 mag. The time scale of the high amplification event would then typically be measured in seconds.

Another possible application of the parallax effect would be to make use of the annual motion of the Earth. For an Earth-bound observer the distance $d(t)$ to the critical curve in the observer plane varies (1) due to the transverse velocity of the solar system relative to the critical curve and (2) due to the annual motion of the Earth. We note that the position of the Earth projected onto the plane of the observer is determined by the celestial coordinates of the observed object.

Figure 2 shows a simulated light curve for the lens object 2237+0305, which because of the large value of D_{dq}/D_d (~10) gives the best chance of observing the effect. Here we have used a dimensionless surface density of continuously-distributed matter of $\sigma_c = 0.8$, a Hubble parameter of $H_0 = 75\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and a standard cosmological model ($q_0 = \sigma_0 = 0$). The source has a radius $R = 2.7 \times 10^{-4}$ pc and a luminosity profile $I(r) = I_0 (1-(r/R)^2)^{1/2}$. This seems to be more reasonable than $I(r) =$ constant and has the advantage of leading to an analytical solution of the mean amplification $\langle A(\mathrm{d})\rangle$. The change in brightness is given by $m(\mathrm{d}) = -2.5 \log \langle A(\mathrm{d})\rangle$. We further assume a transverse velocity of one source radius per year, that is $260\ \mathrm{km\ s^{-1}}$ in the source plane or $25\ \mathrm{km\ s^{-1}}$ in the observer plane. The annual motion of the Earth shows up clearly in the light curve as an oscillation (Fig. 2). It is clear that observable effects only occur if (1) the transverse velocity of the solar system relative to the critical curve is smaller than or about equal to the orbital velocity of the Earth ($v_T \leq 50\ \mathrm{km\ s^{-1}}$ and (2) the projection $\tilde{R}$ of the source radius R onto the observer plane is not too great ($\tilde{R} \leq 100$ AU). The first condition is particularly unlikely to be fulfilled, because a typical transverse velocity in the observer plane is of the order of $500\ \mathrm{km\ s^{-1}}$. It therefore seems safe in most cases to neglect the annual parallax effect by gravitational micro-lensing, even during the high amplification events.

This work was supported in part by the Deutsche Forschungsgemeinschaft under Az. Re 439.3.

Received 11 June; accepted 14 August 1986.

1. Chang, K. & Refsdal, S. *Nature* **282**, 561–567 (1979).
2. Kayser, R., Refsdal, S. & Stabell, R. *Astron. Astrophys.* **166**, 36–52 (1986).
3. Gott, J. R. *Astrophys. J.* **243**, 140–146 (1981).
4. Chang, K. Ph.D thesis, Univ. Hamburg (1981).
5. Chang, K. & Refsdal, S. *Astron. Astrophys.* **132**, 168–178 (1984).
6. Young, P. *Astrophys. J.* **244**, 756–767 (1981).
7. Paczynski, B. *Astrophys. J.* **301**, 503–516 (1986).
8. Chang, K. *Astron. Astrophys.* **130**, 157–161 (1984).
9. Bartels, G. G. thesis, Univ. Hamburg (1981).
10. Schneider, P. & Weiss, A. *Astron. Astrophys.* (submitted).

Companion star of a binary millisecond pulsar

G. A. Wright & E. D. Loh

Department of Physics and Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544, USA

Binary pulsars offer important clues about the evolution of neutron stars. We have searched for and identified the companion star of the recently discovered[1] binary millisecond pulsar PSR1855+09. The optical counterpart is compatible with a $0.3M_\odot$ white dwarf with a surface temperature of ~5,900 K. This is in accord with currently accepted ideas on the genesis of such a fast binary pulsar[2]. The low temperature of the white dwarf implies that the system is $>10^9$ yr old. Such an old pulsar is not easily accounted for by the conventional theory, in which the pulsar ceases to be active after its magnetic field decays, a process believed to take ~10^7 yr.

In February 1986, an accurate position was obtained for PSR1855+09 by Fruchter using the Very Large Array (VLA)[1]. Shortly afterwards, we began to look for the companion of this pulsar. Observations were made in April 1986 with the 2.3-m telescope at the Wyoming Infrared Observatory. Pictures were taken using the Princeton charged coupled device (CCD) camera at the $f/2$ prime focus. The field of view is 7 × 10 arc min and the plate scale is 1.23 arc s per pixel. For the observations of PSR1855+09, 500- and 700-nm bandpass filters were used (each ~100-nm wide), corresponding roughly to the V and R bands.

Nine exposures were made over two nights in the 700-nm band. Of these, the three pictures with the best seeing were averaged and the result used for photometry. Seven 500-nm exposures were taken on one night and the best two averaged. All exposures were 300 s long. Sky transparency was good during the observations. The seeing in the best 700-nm frames is 2.5 arc s (full width at half maximum, FWHM) and 3.0 arc s in the 500-nm frames. The inferior seeing in the 500-nm pictures is due in part to strong, gusty winds during those exposures. The telescope was moved slightly between exposures to avoid always putting the candidate star close to any defective pixels on the CCD.

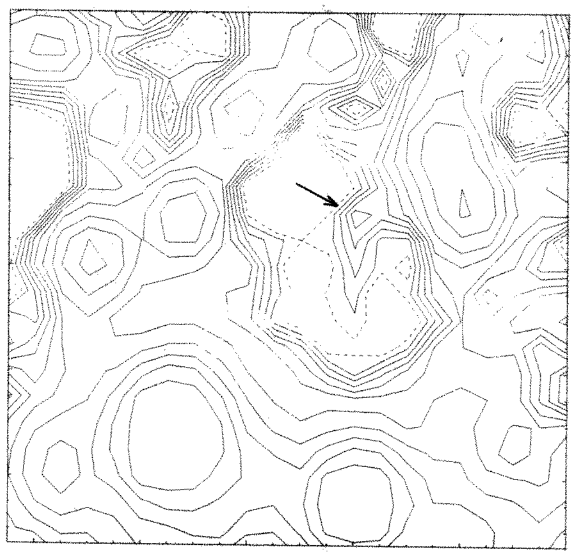


Fig. 1 Isophot plot of the 700-nm-CCD frame of the companion of PSR1855+09. East is up and north is to the right. The arrow indicates the position of the object. The area shown is 27 × 27 arc s. The contours are linearly spaced for low intensities changing to logarithmic at high intensities. The lowest intensity contour is dashed. The west pointing 'tail' near the marked object is caused by a very faint ($m_{700} \simeq 24$) star visible in the best seeing 700-nm frame. The two bright objects at the western edge of the field are $m_{700} \simeq 17.5$ stars visible on Palomar Sky Survey plates.

The candidate star for the companion of PSR1855+09 is shown in Fig. 1. Its peak intensity in the 700-nm frame shown is 7σ above the noise (detector noise plus sky noise). In the poorer seeing 500-nm frame, the peak is 3σ above the noise. The indicated star corresponds, within astrometric error, to the VLA position of PSR1855+09 listed in Table 1. (Error in astrometry is estimated to be ±0.9 arc s in both right ascension and declination.) It is in a dense field of stars, only 3° away from the galactic plane. The number of stars in the field was

Table 1 Parameters of PSR1855+09

Right ascension (1950.0)*	$\alpha = 18$ h 55 min 13.681 s $\pm$ 0.006 s
Declination (1950.0)*	$\delta = +09° 39' 13.39'' \pm 0.09''$
Pulsar period*	$P = 5,362,100,452.5 \pm 0.3$ ps
Orbital period*	$P_b = 1,065,067.49 \pm 0.03$ s
Dispersion measure distance*	$d \simeq 350$ pc
Companion mass*	0.2-0.4$M_\odot$
Pulsar magnetic field*	$B = (3.4 \pm 0.4) \times 10^8$ G
Companion apparent magnitude	$m_{700} = 22.01 \pm 0.10$
	$m_{500} = 22.65 \pm 0.15$
Companion temperature	$T = 5,900^{+1,400}_{-1,000}$ K

* Data from ref. 1.

counted to estimate the probability of a chance coincidence. The measured density of stars gives the probability for the centroid of any star to be within 1 pixel of the indicated position of $\leq 5\%$.

Data reduction and photometry were done using the Princeton Stargaze picture processing system. Photometry is complicated by the fact that the candidate star lies in the wings of a pair of nearby bright stars. These stars were fitted by gaussians and subtracted from the pictures. A 3.7 × 3.7 arc s aperture was centred on the candidate and its intensity measured. A relatively small aperture was chosen to exclude light from other nearby stars. The light falling outside the aperture was estimated by photometry of calibration stars and the intensity was corrected for this. The resulting intensities were converted into AB_{79} system magnitudes using the flux standard stars BD+26 2606 and HD84937 (ref. 3). (AB_{79} magnitudes are relative to a fictitous star whose spectrum is flat in frequency. $m_{AB} = 0$ corresponds to a flux density of 3.65×10^{-20} erg s^{-1} cm^{-2} Hz^{-1}.) The apparent magnitudes are listed in Table 1. Most of the error is due to the uncertainty in the background level of such a crowded field of stars. Other errors come from sky and readout noise and possible errors in centring the photometry aperture. The colour index $m_{500} - m_{700}$ is 0.6 ± 0.2.

The candidate star must be a degenerate star. Main-sequence stars with the observed flux ($m_{700} = 22.0$) in the galactic plane are redder than $m_{500} - m_{700} = 2.3$, because intrinsically blue stars must be distant and highly reddened. Degenerate stars at the same apparent magnitude are ~100 times closer and therefore suffer less reddening. The probability of a degenerate star being with 1 pixel of the pulsar position is $<10^{-4}$ (ref. 4). Therefore we conclude that the star indicated in Fig. 1 is the companion of PSR1855+09.

The temperature of the companion star can be estimated from the colour index. Using the dispersion measure distance obtained by Segelstein *et al.*[1] of 350 pc and taking a value for interstellar absorption of $A_V = 1.9$ mag kpc^{-1} (ref. 5), we get a corrected colour index $m_{500} - m_{700} = 0.3 \pm 0.2$. Dispersion measure distances are typically accurate to 50%, although in some cases they may be off by a factor of two or more[6]. (The error given includes a 50% uncertainty in distance.) The corrected colour index corresponds to a temperature of $5,900^{+1,400}_{-1,000}$ K.

From pulse arrival time measurements, Segelstein *et al.*[1] have estimated the mass of the companion of PSR1855+09 to be 0.2-0.4 solar masses. The observed object is consistent with a white dwarf of $M = 0.3M_\odot$, $R = 0.02R_\odot$, and $T = 5,900$ K at a distance of 475 pc, which agrees with the 350 pc dispersion measure distance.

That the companion star is so cool indicates that the PSR1855+09 system is very old. The characteristic cooling time for a $0.3M_\odot$ white dwarf is 2×10^9 yr (ref. 7). The age of the pulsar, whose progenitor evolves off the main sequence earlier than the white dwarf's, must be greater than this. The pulse period of PSR1855+09 is 5.4 ms. An old pulsar spinning so fast must have been spun up by angular momentum transfer from the companion. Our evidence supports the conventional hypothesis that millisecond pulsars are recycled, that is they are old neutron stars whose large angular momentum is due to mass accreted from a companion star.

A similar low-temperature white dwarf was observed in the PSR0655+64 system by Kulkarni[8]. The existence of such an old yet still active pulsar is at odds with the naive notion of magnetic field decay in pulsars[9], which would switch them off after $\tau_{\text{magnetic}} \simeq 10^7$ yr. Kulkarni proposes that the pulsar magnetic field has two components, one that decays in a timescale of 10^7 yr and a residual nondecaying field. Under some circumstances, the residual field may be sufficient for the pulsar to remain active for much longer than τ_{magnetic}.

Measurement of the pulse period derivative yields an estimate of the surface magnetic field of a pulsar[10]. The timing data[1] for PSR1855+09 gives a field of $(3.4 \pm 0.4) \times 10^8$ G. Because the age of the pulsar is 100 times the field decay time, the present field is a good candidate for a nondecaying remnant field. The other millisecond pulsars, PSR1937+21 and PSR1953+29, have surface fields within a factor of two of this value. These three pulsars have the smallest magnetic fields measured. Whether this represents a lower limit on pulsar magnetic fields is an interesting question awaiting more data.

We thank the director and staff of the Wyoming Infrared Observatory for telescope time and assistance. Earl Spillar helped us make the observations. Andy Fruchter, Lloyd Rawley, Dan Stinebring and Joseph Taylor supplied much useful information and made many helpful comments. This work was supported in part by the NSF.

Received 11 June; accepted 29 August 1986.

1. Segelstein, D. J., Rawley, L. A., Stinebring, D. R. Fruchter, A. S. & Taylor, J. H. *Nature* **322**, 714-717 (1986).
2. van den Heuvel, E. P. J. & Tamm, R. E. *Nature* **309**, 235-237 (1984).
3. Oke, J. B. & Gunn, J. E. *Astrophys. J.* **266**, 713-717 (1983).
4. Liebert, J. *A. Rev. Astr. Astrophys.* **18**, 363-398 (1980).
5. Allen, C. W. *Astrophysical Quantities* 3rd edn, 263 (Athlone, London, 1973).
6. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613-639 (1985).
7. Shapiro, S. L. & Teukolsky, S. A. *Black Holes, White Dwarfs and Neutron Stars*, 86 (Wiley, New York, 1983).
8. Kulkarni, S. R. *Astrophys. J.* **306**, L85-L89 (1986).
9. Bhattacharya, D. & Srinivasan, G. *Curr. Sci.* **55**, 327-330 (1986).
10. Manchester, R. N. & Taylor, J. H. *Pulsars*, 177 (Freeman, San Francisco, 1977).

An explanation for fast downflows on the Sun

A. N. McClymont* & I. J. D. Craig†

* Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA
† Department of Mathematics, University of Waikato, Private Bag, Hamilton, New Zealand

The discovery of coronal loops has provoked major revisions in the theory of solar and stellar atmospheres. The solar corona is now recognized as a highly structured plasma, organized by bundles of intense magnetic field into individual arch- and loop-like structures along which gas can flow[1]. Each arch, rooted in the solar interior, extends through the chromosphere ($T < 10^4$ K), and generally contains coronal plasma ($T > 10^6$ K) in its upper reaches. Early attempts at explaining the temperature and density structure of isolated static loops met with considerable success, despite the simplicity of the theory[2,3]; however, attempts to generalize the theory to describe loops with mass flows have remained singularly unsuccessful[4-9]. Here we point out that the puzzling observational properties of loop mass flows[10-15] can be naturally explained if the loop is sufficiently cool ($T \lesssim 10^6$ K)—that is, if it contains no coronal plasma.

Observed flows exhibit several features that must be explained by any convincing theory. First, the flows are observed to persist much longer than the transit time of a gas element through the

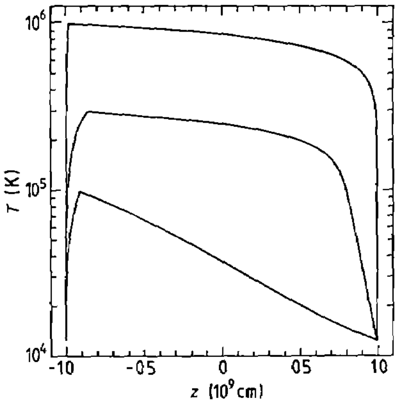


Fig. 1 Temperature profile in a loop of length 2×10^9 cm (z is the distance along the loop) with a steady flow of 5 km s^{-1} at 10^5 K, for peak temperatures of 10^5, 3×10^5 and 10^6 K. The corona is heated by a point energy source (at the position of maximum temperature). Extreme asymmetry is required to drive flows of the observed magnitude, 5–10 km s^{-1} at 10^5 K.

loop[10,11]. This fact alone suggests a quasi-steady model in which plasma flows continually from one footpoint to the other. This model also readily accounts for the second observed property, namely, that emission from the downflowing leg predominates[10-15]. This can be understood as a consequence of the asymmetry of the temperature profile[9] induced by steady flow (see Fig. 1). Altogether more tantalizing, however, is the third property: flow speeds are large (5–10 km s^{-1}) at transition-region temperatures ($T\simeq10^5$ K), yet the flows apparently disappear at coronal temperatures[11-14]. This feature appears completely inexplicable in terms of current steady flow models. The problem arises because, on the assumption of subsonic flow, the pressure $P=2\rho RT$ should be nearly uniform, which, together with the steady-state condition of constant mass flux $q=\rho v$, implies $v\propto T$. It follows that in typical loops the highest velocities are to be expected at coronal temperatures, namely $T\geq10^6$ K.

The key to this paradox lies in momentum conservation. When the flow velocity approaches the sound speed, the dynamic pressure, $\rho v^2\equiv qv$, becomes important, forcing the gas pressure to decrease. The net result is an upper limit on the coronal temperature for a given flow speed (see equation (3) below).

To illustrate the problem in the simplest possible way, we neglect gravity and assume a loop of uniform area. Then the steady flow momentum equation implies

$$qv+P=qv_0+P_0 \tag{1}$$

where subscript zero denotes conditions at a reference point. As the dynamic pressure qv increases, the gas pressure P must decrease. Eliminating $\rho=q/v$ from equation (1) and writing $M=v/c$ for the Mach number relative to the isothermal sound speed $c=\sqrt{2RT}$, we can solve for temperature, obtaining

$$\frac{T}{T_0}=\frac{v}{v_0}\left[(1+M_0^2)-M_0^2\frac{v}{v_0}\right] \tag{2}$$

For subsonic flow ($v\ll c$), equation (2) reduces to the familiar hydrostatic condition of uniform pressure, $v/v_0=T/T_0$. But at higher flow speeds the function $T(v)$ is a parabola, with a turning point at (v^*, T^*), where

$$\frac{v^*}{v_0}=\left(\frac{1+M_0^2}{2M_0^2}\right),\qquad \frac{T^*}{T_0}=\left(\frac{1+M_0^2}{2M_0}\right)^2 \tag{3}$$

which occurs when the flow speed reaches the isothermal sound speed.

This turning point defines the maximum possible temperature T^* in the loop. Setting $T_0=10^5$ K and noting that for $v_0\lesssim10$ km s^{-1}, $M_0^2\ll1$, we obtain from equation (3)

$$T^*\simeq(20/v_0)^2\times10^5\ \mathrm{K} \tag{4}$$

with v_0 in km s^{-1}. It is clear that higher loop temperatures correspond to lower flow speeds. For v_0 in the range 5–10 km s^{-1}, equation (4) yields $T\lesssim4\times10^5$–1.6×10^6 K. Thus the absence of significant flow velocities at coronal temperatures is a natural consequence of momentum conservation.

So far the cause of mass flow has not been considered. A complete description of the flow requires consideration of the energy equation[9,16-18]; in the present model, spatial asymmetry of the coronal energy source powers the flow. Figure 1 shows temperature profiles for peak temperatures of 10^5, 3×10^5 and 10^6 K, with a flow speed of 5 km s^{-1} at 10^5 K, computed for the simplest case of a localized 'point' energy source. The heating asymmetry is immediately apparent. Indeed, extreme heating asymmetry is required to drive flows of the observed magnitude. The flow originates as follows[9]. Energy is transported most readily (by thermal conduction) towards the closer footpoint, where the excess heat flux 'boils off' chromospheric material. The conductive flux into the downflowing leg of the loop is weaker, allowing the gas to cool as it flows towards the distant footpoint, where it condenses back into the chromosphere. Note that the coronal heat source is only ~10% of the loop length above the chromosphere in the upflowing leg. Larger flow speeds require even greater asymmetries[18].

Turning briefly to the question of gravity and area non-uniformity, we find that these can significantly alter the flow pattern, but in opposite senses. Gravity causes the gas density to decrease with increasing height, thus the velocity of upflowing plasma, $v=q/\rho$, tends to increase more quickly, enhancing the influence of dynamic pressure and reducing the maximum temperature (equation (3)) still further. Fast flows may therefore be constrained not only to cool loops, but also to low-lying loops. If the loop area A increases with height, the velocity, $v=q/A\rho$, rises less quickly, so that the constraints on loop temperature and height are somewhat relaxed. However, physically realistic area variations seem unlikely to produce a major effect[18].

Thus we conclude that all features of the observed flows follow from the analysis presented here. Specifically, steady flows that exhibit substantial transition region velocities with a predominance of redshifted emission, and which disappear at coronal temperatures, are naturally explained. The theory will be strengthened if spatially resolved observations can confirm the predicted asymmetry in loop structure associated with fast flows.

This research was partially supported by NSF grant ATM85-96006.

Received 24 March; accepted 19 June 1986.

1. Withbroe, G. L. in *Solar Active Regions* (ed. Orrall, F. Q.) 199–212 (Colorado Associated University Press, Boulder, 1980).
2. Craig, I. J. D., McClymont, A. N. & Underwood, J. H. *Astr. Astrophys.* **70**, 1–11 (1978).
3. Rosner, R., Tucker, W. H. & Vaiana, G. S. *Astrophys. J.* **220**, 643–655 (1978).
4. Yeh, T. *Sol. Phys.* **55**, 241–250 (1977).
5. Pneuman, G. W. & Kopp, R. A. *Sol. Phys.* **57**, 49–64 (1978).
6. Cargill, P. J. & Priest, E. R. *Sol. Phys.* **65**, 251–269 (1980).
7. Athay, R. G. *Astrophys. J.* **249**, 340–348 (1981).
8. Noci, G. *Sol. Phys.* **69**, 63–76 (1981).
9. Craig, I. J. D. & McClymont, A. N. *Astrophys. J.* **307**, 367–380 (1986).
10. Gebbie, K. B. *et al.* *Astrophys. J.* **251**, L115–L118 (1981).
11. Feldman, U., Cohen, L. & Doschek, G. A. *Astrophys. J.* **255**, 325–328 (1982).
12. Dere, K. P. *Sol. Phys.* **77**, 77–93 (1982).
13. Doschek, G. A., Feldman, U. & Bohlin, J. D. *Astrophys. J.* **205**, L177–L180 (1976).
14. Lites, B. W. *et al.* *Astrophys. J.* **210**, L111–L113 (1976).
15. Athay, R. G., Gurman, J. B., Henze, W. & Shine, R. A. *Astrophys. J.* **265**, 519–529 (1983).
16. Boris, J. P. & Mariska, J. T. *Astrophys. J.* **258**, L49–L52 (1982).
17. Mariska, J. T. & Boris, J. P. *Astrophys. J.* **267**, 409–420 (1983).
18. McClymont, A. N. & Craig, I. J. D. *Astrophys. J.* (in the press).

Poly(styrene peroxide): an auto-combustible polymer fuel

K. Kishore & T. Mukundan

Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bangalore 560 012, India

Commonly, polymers exhibit extremely slow, candle-like burning, although oxidizers can increase the rate of combustion, as in solid propellants. Polymer degradation occurs before combustion, and being endothermic, demands a continuous supply of energy. However, if a polymer can degrade exothermically, it can sustain degradation and automatically increase the rate of combustion (autopyrolysis). Poly(styrene peroxide), PSP, which degrades exothermically[1] is such an exceptional and rare example. We explored the possibility of autopyrolytic behaviour in PSP. Interestingly, it was observed that, once initiated, PSP not only sustains its own degradation, but also supports a flame in air, with a burning rate comparable to that of solid propellants. This makes PSP an ideal candidate for a special fuel, such a fast-burning autopyrolytic polymer being hitherto unknown. Here, in the first reported study on the combustion of a polymeric peroxide, we examine some salient features of the combustion of this unique auto-combustible polymer and potential fuel, and show that the combustion is controlled by polymer degradation.

PSP was prepared and purified from styrene and oxygen as reported by Miller and Mayo[2]. It was characterized by ^1H and ^{13}C NMR, IR and Raman spectroscopy. The PSP samples containing additives (nitrobenzene and *p*-toluidine) were prepared by dissolving PSP and the additive (one mole for every 20 polymer repeat units) in acetone, followed by removal of solvent. The polymer was filled free of air bubbles in glass tubes 8 mm in diameter with cotton support. The sample was ignited using a nichrome wire, and the linear burning rate was measured. The burning rates (the average of five repititions) were closely reproducible (within 1% error). Temperature profile measurements were carried out using a chromel-alumel thermocouple (0.127 mm in diameter) inserted into holes drilled into the polymer sample so that the thermocouple bead was well inside. The thermocouple output was fed to a Philips (PM 3234, 0-10 MHz) double beam cathode ray storage oscilloscope, maintained at a sensitivity of 5.8 mV cm^{-1}. The sample tubes were insulated with plaster of paris to minimize lateral heat losses. The degradation peak temperatures of PSP with and without additives were obtained on a DuPont 990 differential scanning calorimeter at a heating rate of 20 °C per min with a sample weight of ~2 mg. The temperature axis was calibrated using benzoic acid.

Pyrolysis gas chromatography studies of the same samples were carried out on a Shimadzu GC-4C gas chromatograph, with attached PYR-1A pyrolyser using a dioctyl phthalate coated on a chromosorb column. Nitrogen (35 ml min^{-1}) was used as carrier gas. The samples were pyrolysed at 200 °C, and chromatographed on the column heated at 5 °C min^{-1}, from 50-125 °C. The products were identified by comparison with retention times of pure compounds. The three sample weights varied and hence the aldehyde peak intensities are not the same.

Table 1 Degradation peak temperatures and burning rates of PSP with and without additives

Additives (5 mole % per repeat unit of PSP)	DSC peak temperature (°C)	Linear burning rate (mm s^{-1})
Nil	149.0±0.2	2.00±0.03
p-Toluidine	141.8±0.2	2.47±0.03
Nitrobenzene	158.4±0.2	0.162±0.001

DSC, differential scanning calorimetry

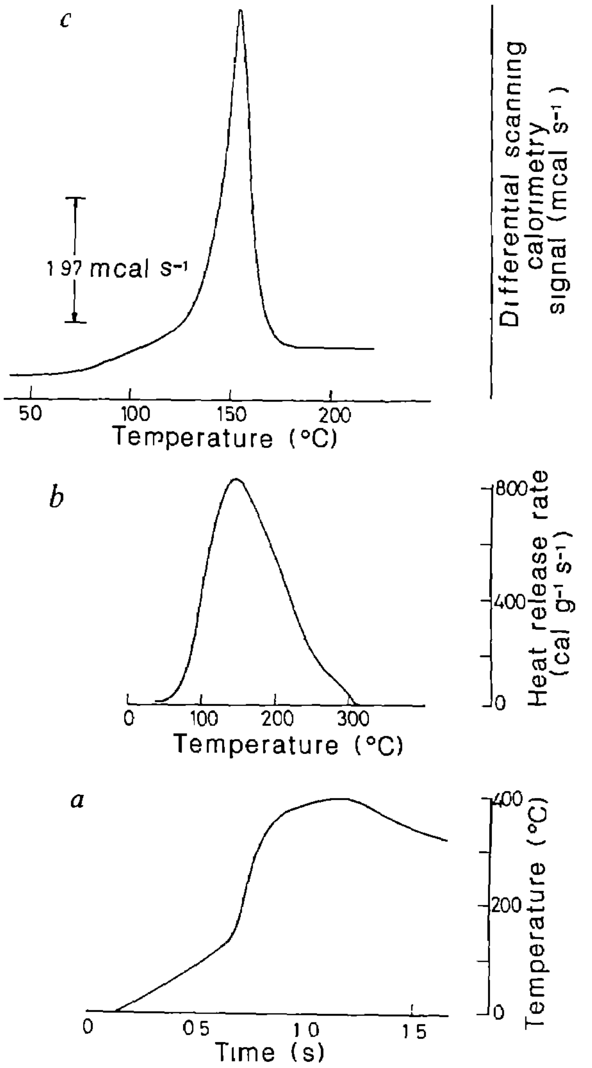


Fig. 1 Thermal profile of poly(styrene peroxide). *a*, Time-temperature trace during combustion, *b*, computer generated, rate of heat release against time plot, *c*, differential scanning calorimetry trace of degradation, heating rate = 20 °C min^{-1}, sample weight ~2 mg

PSP degrades via a radical chain scission mechanism[3] giving benzaldehyde and formaldehyde as major products. A recent pyrolysis gas chromatography study (in the pyrolysis range 120-430 °C) confirms the formation of these two products and accounts for more than 90% of the total degradation products[4]. The remaining products might be formed due to the radical and ionic disproportionation mechanism as proposed by Mayo and Miller[3].

A typical temperature profile (time-temperature trace) is shown in Fig. 1*a*. Heat release rate (dq/dt) at any instant can be calculated from the equation,

$$\frac{dq}{dt} = C_p \frac{T_2 - T_1}{t_2 - t_1}$$

T_2 and T_1 being the temperatures at time t_2 and t_1 respectively, and C_p the average heat capacity of PSP. The profile data were fed into the computer to obtain a plot of dq/dt against T. The temperature dependence of C_p, measured by differential scanning calorimetry was obtained using a computer fit ($C_p = 0.001236T + 0.3541$). The value of C_p above 90 °C, when PSP starts to degrade during combustion, was determined by extrapolation of this fit. The plot of dq/dt against T is given in Fig. 1*b*. A typical calorimetric trace (of PSP degradation) that relates the rate of heat flow (dq/dt) to time/temperature, is shown in

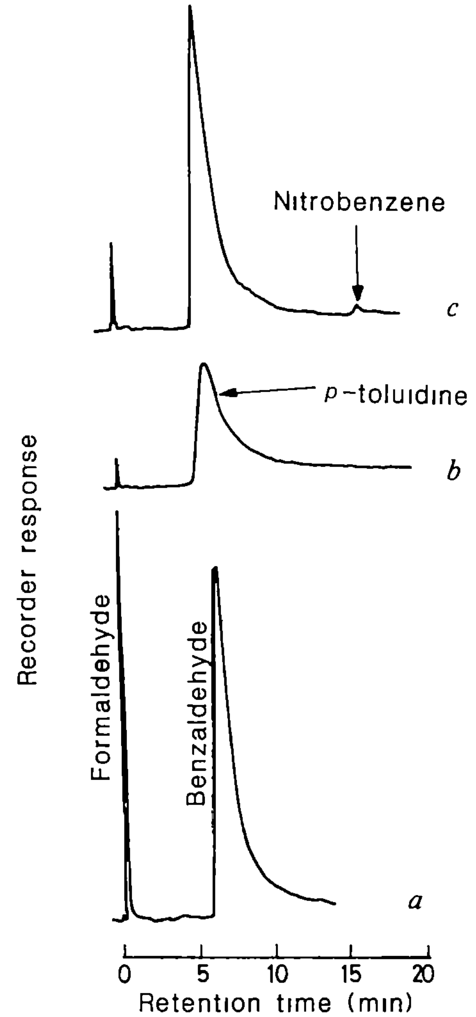


Fig 2 Pyrograms of PSP samples *a*, PSP, *b*, PSP+*p*-toluidine (merges with benzaldehyde) and *c*, PSP+nitrobenzene

Fig 1*c* Note that the solid-gas interface temperatures (surface temperatures) where PSP degradation occurs coincides in the two cases shown in Fig 2*b*, *c*

The profile experiment was used to determine the heat of degradation of PSP during its combustion PSP initially at ambient temperature degrades and the gaseous products are heated to a maximum temperature before entering the flame at the mouth of the tube The enthalpy change for PSP going from ambient to maximum gas temperature is $\Delta H = C_p(T_2 - T_1)$, T_2 and T_1 being the maximum gas temperature (400 °C) in the profile (Fig 1*a*) and ambient temperature (25 °C) respectively, an average value of $C_p = 0\ 54$ cal mol^{-1} at the surface temperature, 150 °C, was used Corrections for vaporization of benzaldehyde and formaldehyde[1] were made to the above value of ΔH and the heat of degradation of PSP from the combustion experiment thus obtained was 45 0 kcal mol^{-1} Considering the heat losses that occur during the combustion experiment this value is in reasonable agreement with the experimental value reported earlier[1] Comparing the heat of degradation value obtained from the profile experiment and those from differential scanning calorimetry (50±1 kcal mol^{-1}) and thermochemical calculations[1] (52±1 5 kcal mol^{-1}) suggests that combustion of PSP occurs as a consequence of its degradation, otherwise heat released before combustion should differ from the heat of degradation obtained by scanning calorimetry This conclusion is supported by the following observations (1) During combustion, the flame remains at the mouth of the tube, while the surface at which degradation occurs recedes, that is, the flame persists as a result of the combustion of PSP degradation products, (2) under controlled conditions PSP degrades without a flame, the regression rate being equal to the burning rate

A corroboration of this inference was sought with the help of two additives Thermograms were recorded to obtain degradation peak temperatures with and without these additives, the linear burning rates were also measured (Table 1) It can be seen that the additives bring about similar changes in combustion and degradation, *p*-toluidine accelerates both, while nitrobenzene decelerates both

Pyrolysis gas chromatography of the PSP samples containing additives was carried out to investigate the effect on PSP degradation (Fig 2) No new degradation products were formed in the presence of the additives and the overall composition of the degradation products remained the same In the case of the PSP-nitrobenzene sample, nitrobenzene has a longer retention time (Fig 2*c*), while for the PSP-*p*-toluidine case the *p*-toluidine merges with the benzaldehyde peak (Fig 2*b*) These results suggest that the additives only affect the rate of reaction rather than the mechanism of PSP degradation A detailed study of this aspect, which unavoidably falls beyond the realm of this preliminary report, would prove vital in regulating the rate of combustion for potential use of the polymer as a fuel However, the present investigation describes the unusual phenomena of autopyrolysis and autocombustion of an unadulterated polymer and unambiguously proves that its combustion is primarily controlled by degradation

We thank Mr P A Joy and Mr M Jayamurthy for their help in computations

Received 17 March, accepted 5 September 1986

1 Kishore, K & Ravindran, K *Macromolecules* **15**, 1638-1639 (1982)
2 Miller, A A & Mayo, F R *J Am Chem Soc* **78**, 1017-1023 (1956)
3 Mayo, F R & Miller, A A *J Am. Chem. Soc* **78**, 1023-1034 (1956)
4 Kishore, K & Ravindran, K *J analyt appl Pyrol* **5**, 363-370 (1983)

Hydrothermal methane plumes in the Mariana back-arc spreading centre

Y. Horibe*‡, K.-R. Kim†‡ & H. Craig†

* Ocean Research Institute, University of Tokyo, Nakano,
Tokyo 164, Japan
† Scripps Institution of Oceanography, University of California at
San Diego, La Jolla, California 92093, USA

Submarine hydrothermal vents were first unequivocally detected by measurement of ^{3}He/^{4}He ratio anomalies coupled with deviations from ambient T–S relationships on the Galapagos spreading centre[1,2]. Since then the existence of hydrothermal vents along the entire length of the East Pacific Rise has been demonstrated by shipboard observations of ^{3}He, methane, and manganese in the effluent plumes[3–5]. We have extended our studies of vent fields at mid-ocean-ridge spreading centres to search for effluent plumes in back-arc basins, which are also thought to be active regions of seafloor spreading where new oceanic crust is formed[6], and thus where rifting and fissuring along the axis of extension should provide similar opportunities for the downward penetration of sea water and the development of active hydrothermal systems. We have detected large plumes of methane-enriched water in the Mariana Trough back-arc basin[7] and also in the summit crater of Loihi Seamount[7] at the present site of the Hawaiian hotspot, thus establishing the widespread existence of 'off-ridge' hydrothermal vents in the ocean. However, although the Loihi vents produce plumes enriched in both ^{3}He and methane[7], we find that, surprisingly, the Mariana vents are enriched in methane without a corresponding enrichment in ^{3}He This is the first time such a 'decoupling' of these two effluent plume tracers has been observed.

---

‡ Present addresses Scripps Institution of Oceanography, La Jolla, California 92093, USA (Y H), Department of Oceanography, Seoul National University, Seoul 151, Korea (K -R K)

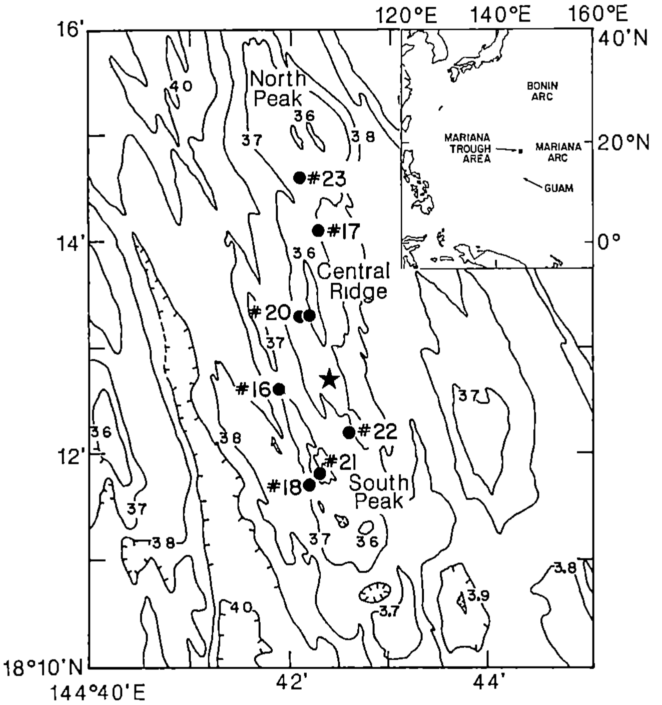


Fig. 1 Cepheus station locations in the Mariana Trough The stations lie between two large volcanic peaks four miles apart, along a ridge of fresh lavas from fissure eruptions This section is a high spot in the Mariana Trough and was chosen for investigation because the axis is accessible for study by the submersible Alvin (depth < 4 km) The star on Central Ridge marks the site of a 0 02 °C temperature anomaly observed with Deep Tow by P Lonsdale, who made this map A CTD towed on the Cepheus expedition also detected a 'true' hydrothermal plume spike in the vicinity of station 20, with amplitude 0 01 °C Depth contours in km

Our search for off-ridge hydrothermal vents was carried out in a region of the axial rift valley between 18°11′ N to 18′15′ N, at the suggestion of P Lonsdale, who had previously (with the Deep Tow CTD (conductivity, temperature, depth) observed small temperature spikes in the water column that were 'real' plumes in the sense that corresponding decreases in salinity were lacking[1] The shipboard work was done on the Japanese vessel *Hakuho Maru* on the Cepheus expedition[8], in the area shown in Fig 1, which was mapped by Lonsdale Seven hydrographic stations were occupied along the axial volcanic ridge, a methane plume was observed in all cases although complete profiles through the plume were obtained at only four of the stations

Figure 2 shows the Mariana methane plume in section from south to north along the ridge crest The core concentration of methane is 74×10^{-6} cm^3 kg^{-1} at a depth of 2950–3000 m, increasing slightly toward the northern peak The plume rises 600–700 m above the bottom before entraining enough ambient water to reach density equilibrium and spread horizontally along the 3000 m isopycnal (1041 428 kg m^{-3} at 2957 m) This 'upper' plume is distinctly separated from the lower more discontinuous plume ~200 m above the sea floor, and it is obviously generated by more active vents We suppose that the 'upper' plume is a '21° N'-type plume generated by the so-called 'black smokers', which vent high-temperature water from chimneys above the sea floor, whereas the lower and much more dilute plumes probably represent the low-temperature 'Galapagos-type' effluent from a cooled subsurface reservoir These lower plumes are discontinuous in section because of the offset of the axial ridge at 18°12 5′ N (Fig 1) We have observed[9] similar sets of near-bottom and 'upper' plumes on the East Pacific Rise at 20° S, where the plumes are clearly associated with vents as shown by Angus photographs of the vent fauna However, the East Pacific Rise plumes generally rise only some 300 m above the ridge crest before spreading horizontally[9,10], much less than the height observed here As we observed that the potential density gradients are quite similar both in the Mariana plume and at several localities on the East Pacific Rise, it is possible that the much greater rise height of the Mariana plume is due to higher effluent temperatures, perhaps above 400 °C, from very large 'black smokers'

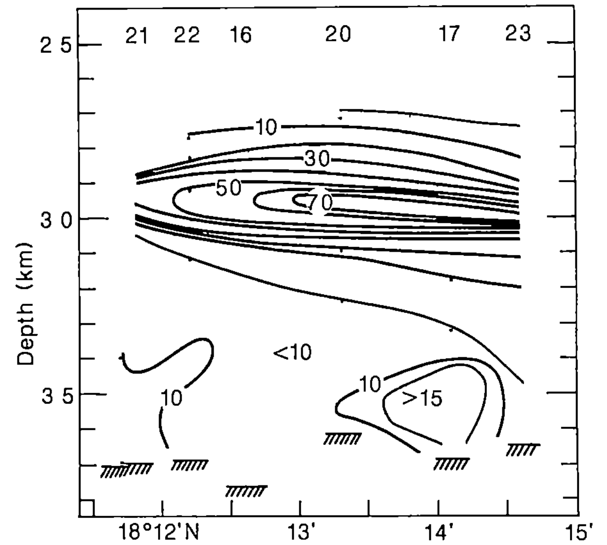


Fig 2 Structure of the methane plumes along the Mariana Trough spreading axis Station numbers are along the top, methane contours are in 10^{-6} cm^3 (STP) methane per kg sea water Bottom depths at the stations are indicated by cross-hatched bars

Table 1 shows the methane and helium data at station 23, the only station for which a helium profile was obtained The ^{3}He/^{4}He ratio anomalies are tabulated as $\delta\,^3$He values (in per cent), that is, the percentage enrichments relative to the atmospheric ratio The saturation anomalies relative to atmospheric solubilities[11] are also listed for both He isotopes and for neon, and finally the absolute He isotope concentrations are tabulated for each isotope

Figure 3 shows the $\delta\,^3$He profile superimposed on the methane plumes from stations 20 and 23, and it is seen that the very large and sharp methane maximum is not reflected in the ^{3}He profile A close reading of Table 1 shows that there is probably a very slight maximum in $\delta\,^3$He between 3,000 and 3,145 m, correlated with a similar slight maximum in the solubility anomalies and absolute concentrations of the He isotopes The solubility anomalies and absolute concentrations of ^{3}He indicate that there may be a small maximum at 3,000 m, amounting to no more than 0 8% Comparison with the methane peak indicates that the methane/^{3}He ratio in the added volatiles must be $> 100 \times 10^{-6}$, which is much greater than the $\sim 3 \times 10^{-6}$ observed at East Pacific Rise vent fields[5,12] We also find that, in detail, the ^{3}He and ^{4}He concentrations are well correlated ($r = 0$ 87), with a slope corresponding to a mixture of a He component with a ^{3}He/^{4}He ratio of only 3 2 (± 0 7) times atmospheric, rather than a ratio of 8 as observed on mid-ocean ridges This indicates that there is a significant radiogenic He component in the Mariana vents in addition to the large methane enrichment relative to He These data as a whole thus suggest that some effects we have not encountered on the East Pacific Rise have significantly influenced the volatile compositions in the Mariana vents

A methane profile was also measured in the so-called 'Mounds' area west of the axis of the Mariana Trough at 18°01′ N, 144°18′ E, an area of high heat flow thought to be a site of hydrothermal activity[13] Samples were collected at 12 depths from 2,930 to 3,515 m, above the bottom depth of 3,534 m The observed methane concentrations had a mean of (4 3 ± 0 4) × 10^{-6} cm^3 kg^{-1}, which is identical to the ambient methane back-

Table 1 Methane, ^{3}He, ^{4}He, Ne and potential temperature (T_{pot}) measurements for Cepheus expedition station 23

Depth (m)	T_{pot} (°C)	Methane (10^{-6} cm^3 kg^{-1})	δ^3He (%)	ΔNe (%)	Δ^4He (%)	Δ^3He (%)	^{4}He (10^{-6} cm^3 kg^{-1})	^{3}He (10^{-12} cm^3 kg^{-1})
2,705	1 520	4 6	—	—	—	—	—	—
2,803	1 501	8 5	25 9	(3 60)	7 54	37 1	42 46	74 9
2,852	1 486	14 7	27 2	(4 12)	7 99	39 0	42 64	75 9
2,901	1 472	19 4	26 7	(5 18)	7 70	38 1	42 53	75 4
2,949	1 487	45 0	27 4	3 30	7 89	39 1	42 60	76 0
2,998	1 482	71 7	27 6	3 18	8 06	39 5	42 67	76 2
3,047	1 467	33 0	27 7	3 19	8 11	39 7	42 69	76 3
3,145	1 457	16 6	27 8	(3 52)	7 73	39 4	42 54	76 1
3,242	1 457	8 5	27 3	(3 83)	8 12	39 3	42 70	76 1
3,340	1 456	7 9	27 3	3 09	7 80	38 9	42 57	75 9
3,438	1 436	3 4	26 7	3 07	7 86	38 3	42 60	75 6
3,536	1 435	3 5	27 3	3 19	7 99	39 1	42 65	76 0

Seawater samples were collected in 23-litre Niskin hydrographic bottles, with the use of a pinger to measure distance above the sea floor Samples of 40 ml (He) and 1 liter (methane) were drawn in tandem in copper tubing and Pyrex cylinders The He samples were sealed with refrigeration clamps and were extracted at Scripps Institution of Oceanography and analysed for ^{3}He/^{4}He ratios and He and Ne concentrations (by isotope dilution) as described by Lupton and Craig[10] Methane measurements were made on board with the Shimadzu FID (flame ionization detection) gas chromatograph and vacuum extraction system described by Kim[9]

The salinity at 2,705 m is 34 657‰, below this depth the salinity is (34 663 ± 0 002)‰, that is, constant to within the precision of the measurements[8] δ^3He $= (R/R_A - 1)\times 100$, where $R = {}^3$He/^{4}He in the sample and R_A refers to atmospheric ^{3}He/^{4}He ratio ($=1\ 40\times 10^{-6}$) The Δ values for Ne, ^{3}He and ^{4}He are the saturation anomalies relative to atmospheric solubility at the potential temperature of the sample[11] Slight corrections for air leakage into the copper samplers have been made to the He results for five of the samples in which excess Ne was observed (ΔNe values in parentheses) ΔNe values between 3 0 and 3 3% are normally distributed owing to random errors, and the mean of these values, 3 17%, is taken as the actual Ne solubility anomaly For the five samples with ΔNe > 3 3%, the ΔHe values are reduced by subtracting 1 276 (the solubility coefficient ratio) times the Ne solubility deviation from the mean The recalculated anomalies in ^{3}He and ^{4}He ratios and solubilities, and the absolute concentrations of ^{3}He and ^{4}He, are tabulated for these five samples The precision of the ^{4}He concentration data is ~0 3% The precision of the δ^3He measurements is about 0 7%, so that the corrections for air leakage for these data are not significant except for a 0 6% correction for the 2,901 m sample However, these corrections were made very carefully because the lack of a sharp maximum in the isotope ratios or absolute concentrations made us look very closely at the data

ground in the Mariana Trough samples, (3 9 ± 0 6) × 10^{-6} cm^3 kg^{-1}, obtained by averaging 23 samples with minimum methane concentrations in the same depth range at stations 16–23 on the axis of the trough Thus if hydrothermal activity is present at the 'Mounds' site, it is not intense enough to make observable methane plumes and it must be extremely diffuse

In conclusion, we have observed a large hydrothermal methane plume rising 700–800 m above the sea floor and extending several miles along the axial region of the Mariana Trough The discovery of this plume extends the source of submarine hydrothermal vents to the extensional region of back-arc basins, which implies that there is much more hydrothermal activity in the ocean than previously thought However, the lack of a strong helium signal associated with the methane plume, and the relatively low ^{3}He/^{4}He ratio in the added He, indicate that significant differences in chemistry are present in the mantle sources and hydrothermal water-rock interaction processes in the marginal basins of the oceans Thus we look to the forthcoming Alvin investigations in these regions to increase our understanding of these effects

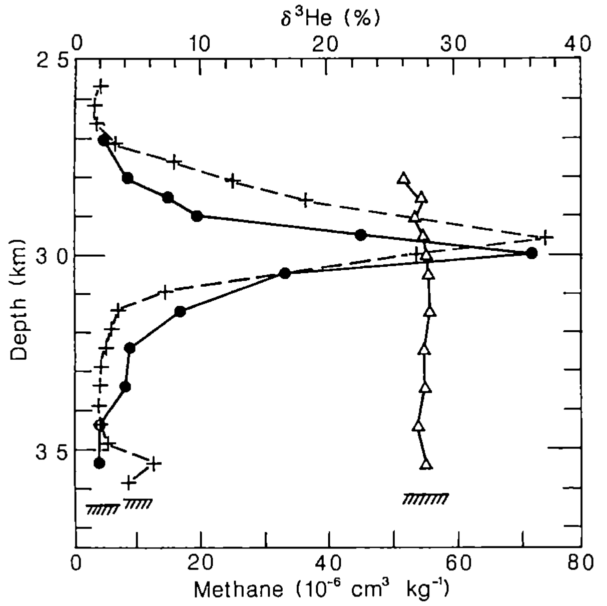


Fig 3 Methane plumes at stations 20 (dashes) and 23 (solid), compared with the δ^3He isotopic ratio anomaly relative to atmospheric helium at station 23 (triangles)

We thank H Tsubota, M Fukazawa, T Gamo, H Inoguchi, H Morinaga and the officers and crew of the vessel for assistance aboard the *Hakuho Maru* We are grateful to Peter Lonsdale for his detailed map and discussions of the Mariana Trough area, and to R L Chase for helpful comments The work at La Jolla was supported by the NSF Division of Ocean Sciences

Note added in proof Since this paper was submitted we have extended our observations of back-arc hydrothermal methane plumes to the central spreading axis of the Lau Basin in the southwest Pacific[14] Hawkins and Helu[15] have dredged sulphides of probable hydrothermal origin from this basin, and Both *et al*[16] have photographed hydrothermal chimneys and vent fauna on the spreading axis of the Manus Basin, where we subsequently observed large methane plumes (Papatua expedition, 1986)

Received 6 May, accepted 13 August 1986

1 Weiss, R F, Lonsdale, P, Lupton, J E, Bainbridge, A E & Craig, H *Nature* **267**, 600-603 (1977)
2 Lupton, J E, Weiss, R F & Craig, H *Nature* **267**, 603-604 (1977)
3 Lupton, J E *et al* *Earth planet Sci Lett* **50**, 115-127 (1980)
4 Craig, H *Eos* **62**, 893 (1981)
5 Kim, K-R, Craig, H & Horibe, Y *Eos* **64**, 724 (1983)
6 Kang, D E *J geophys Res* **76**, 2542-2561 (1971)
7 Horibe, Y, Kim, K-R & Craig, H *Eos* **64**, 724 (1983)
8 Horibe, Y Prelim *Rep CEPHEUS Exped 1982*, 1-78 (Ocean Research Inst, Tokyo, 1983)
9 Kim, K-R Ph D thesis, Univ California at San Diego (1983)
10 Lupton, J E & Craig, H *Science* **214**, 13-18 (1981)
11 Weiss, R F *J chem Engng Data* **16**, 235-241 (1971)
12 Welhan, J A & Craig, H *Geophys Res Lett* **6**, 829-831 (1979)
13 Lonsdale, P & Hawkins, J *Bull geol Soc Am* **96**, 940-951 (1985)
14 Craig, H, Kim, K-R & Poreda, R *Eos* **67**, 378 (1986)
15 Hawkins, J & Helu, S *Eos* **67**, 378 (1986)
16 Both, R *et al* *Eos* **67**, 489-490 (1986)

Structure of the Kerguelen Plateau province from Seasat altimetry and seismic reflection data

Millard F. Coffin*, Hugh L. Davies* & William F. Haxby†

* Bureau of Mineral Resources, Geology and Geophysics,
GPO Box 378, Canberra, A C T 2601, Australia
† Lamont-Doherty Geological Observatory, Palisades,
New York 10964, USA

The Kerguelen Plateau in the southern Indian Ocean stands 2–4 km above the adjacent sea floor and is ~2,500 km long (Fig 1)[1,2]. Seasat provided a unique data set for deriving the free-air gravity field of the region (Fig. 2)[3–5]. Here we report the results of a new analysis of the plateau province's structure employing both Seasat and newly-acquired multichannel seismic (MCS) data[6,7] for ground truth. The northern sector is characterized by volcanism and a sedimentary basin; the southern sector by a broad anticlinal arch, major faulting, and a sedimentary basin; and the eastern sector by an abyssal basin (Labuan) and bounding ridge (William's). The three sectors argue for a more complex tectonic evolution of the feature than has been previously proposed.

Seasat's radar altimeter, pulsing 1,000 times per second, allowed the determination of the average distance between the satellite and several square kilometres of the sea surface directly beneath it One-second averages of the distances, each representing ~8 km of ground track, provided continuous profiles of sea surface topography to a precision of 5–10 cm The ocean surface closely coincides with the geoid, the gravitational equipotential surface reflecting the mass distribution within the Earth Variations in gravitational acceleration computed directly from the geoid surface[4] result in gravity anomaly estimates accurate to 5 mGal with a spatial resolution of 25–50 km Although less accurate than shipboard gravity measurements at spatial scales <100 km, Seasat's uniform track coverage (<5 to several tens of kilometres spacing)[5] permits comprehensive mapping of the gravity field with consistent accuracy (Fig 2), whereas shipboard measurements are sparse in the plateau region (Fig 1) and are inadequate for regional analyses Caution must be exercised in viewing the gravity map south of ~60 °S sea ice increased altimeter noise by an order of magnitude and thus reduced precision, features of dimensions less than several tens of kilometres may be artefacts of the sea-surface gridding procedure or represent measurement noise, and features lineated in the satellite track directions may be artificial, resulting from uncertainties in correlating measurements along adjacent profiles

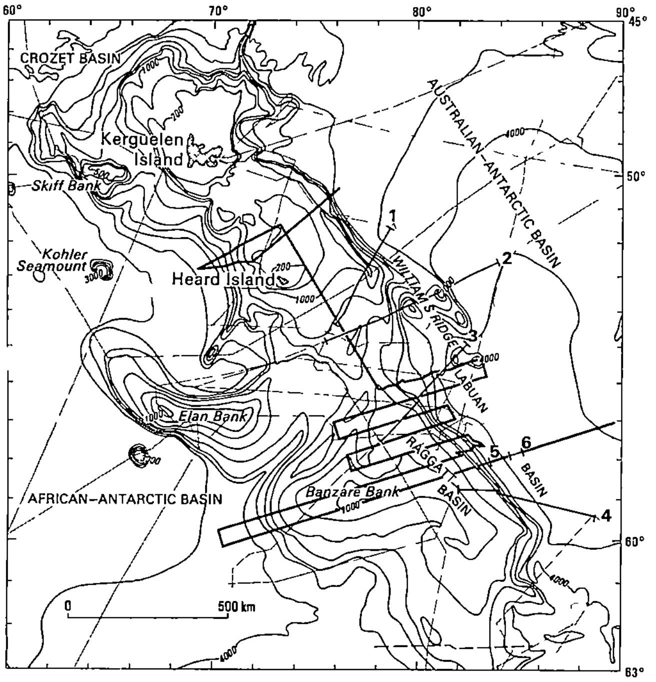


Fig 1 Bathymetry of the Kerguelen Plateau based on shipboard echo-sounding and the Seasat-derived free air gravity field, contour interval is 500 m Seismic control is indicated as follows thick solid lines, 48-channel seismic, dashed lines, digital single-channel seismic, thin solid lines, analogue single-channel seismic All lines have gravity, magnetics, and bathymetry Numbered segments (1–6) indicate profiles displayed in Fig 4

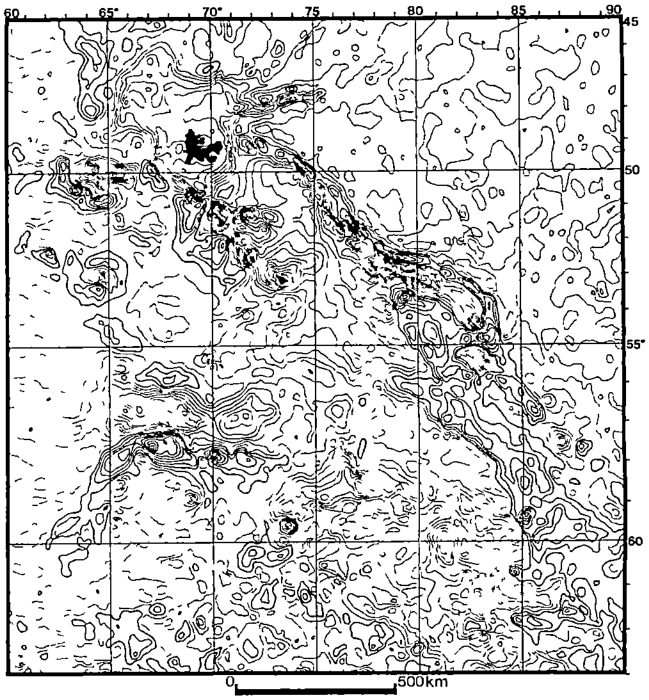


Fig. 2 Seasat-derived free-air gravity field of the Kerguelen Plateau The contour interval is 8 mGal, and the 0 mGal contour is drawn in a heavy black line Kerguelen and Heard islands are in solid black Compare with bathymetry in Fig 1

In 1985, the R/V *Rig Seismic* acquired the first multichannel seismic data over the southern and eastern sectors of the Kerguelen Plateau province[6,7] Two airguns (500 cubic inches) comprised the sound source, and the 1,200-m streamer was configured for 48 channels, allowing 12-fold coverage The record length was 7 s, with a 2 ms sampling rate The data are high quality and reveal far more structural detail than do the older single channel seismic data[1]

Figure 3 summarizes the new structural interpretation of the Kerguelen Plateau province The northern sector includes the province's only islands, composed of volcanics and volcanics/intrusives[8], and offshore comprises a major sedimentary basin[9] as well as numerous smaller basins separated by igneous intrusions or basement horsts[1,2,6,7,9–11] A large positive gravity anomaly coincides with the north-west end of the plateau, defining an apparent structural entity we term the 'Kerguelen massif' The massif, sub-rectangular except for a narrow ridge extending east-northeast from the northeastern margin, includes the Kerguelen Islands in the south and small sedimentary basins in the north and north-east South-southeast of the Kerguelen massif, and separated from it by a transverse gravity trough, is an amalgamation of generally positive gravity anomalies extending as far as Heard Island The largest of the positive anomalies

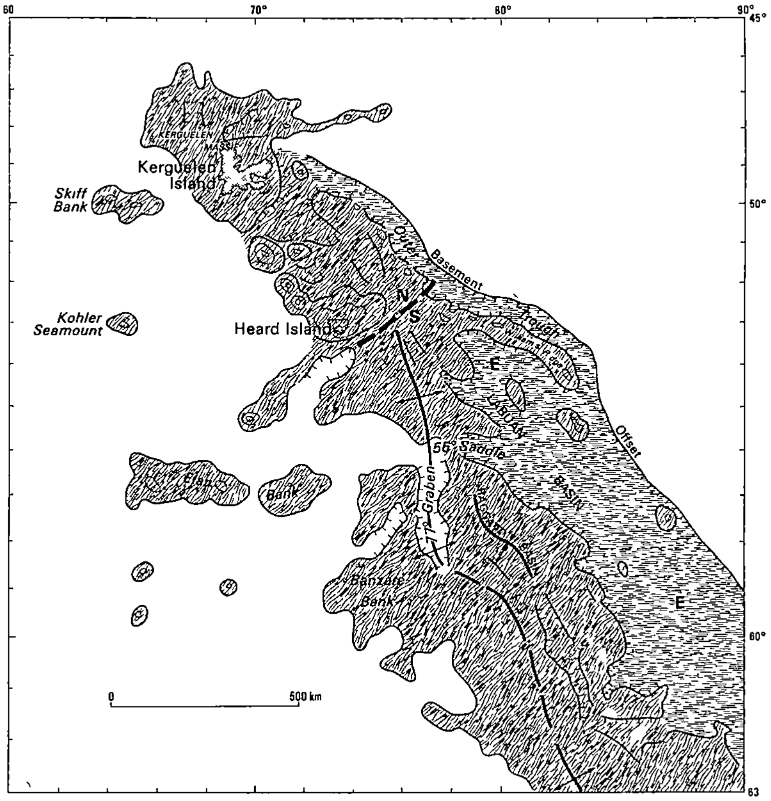


Fig 3 Morphotectonic map of the Kerguelen Plateau province based on combined analysis of Seasat-derived gravity and seismic reflection data Diagonal stippling indicates elevated areas, horizontal stippling indicates deep basins Scarp symbols indicate moderate to steep gravity gradients, in some cases coinciding with normal faults N/S, boundary between northern and southern sectors, E, eastern sector Anticline symbol indicates axis of broad arch of southern sector and syncline symbol indicates axis of Raggatt sedimentary basin The names William's Ridge and Labuan Basin have been formally proposed to the Australian Hydrographer for these newly-defined features

extends for some tens of kilometres north-east and north-west from Heard, and probably demarcates the volcanic pedestal of the island Other discrete, strong sub-circular positive anomalies extend north-west from Heard Island in a rectilinear chain; the major sedimentary basin[9] lies north-east of the chain A series of more subdued positive anomalies extends north-west along the north-east margin of the plateau Where MCS data are available, we observe that the positive anomalies coincide with volcanic pedestals, horsts, or peaks of intrusive rock[6,7] Intervening irregularly-shaped negative anomalies coincide with small, well-developed sedimentary basins. The positive anomalies along the north-east margin, although not obviously horsts, lie along trend from horst blocks of the southern and eastern sectors

The boundary between the northern and southern sectors lies on a transverse line immediately south-east of Heard Island, and is marked by a distinct change in gravity signature and structural style The bulk of the southern plateau is characterized by a more subdued gravity field and a simpler structural regime than the northern sector The southern sector comprises a single broad anticlinal arch, disturbed by normal faults, but, unlike the northern sector, not obviously intruded by igneous rocks Fault motion appears to have occurred over a long time interval[6,7] In places faulting, non-deposition and/or current erosion have resulted in exposure of igneous basement on the sea floor From the boundary with the northern sector at ~53 °S southwards to a prominent saddle at ~56 °S, the anticlinal arch plunges gently south-southeast, it is bounded to the north-east by a series of horsts and grabens[9] but the south-west flank is a simple, gentle slope. Coincident with the saddle at 56 °S, the arch may be cut by a transverse graben South-southeast of the saddle the arch rises and reaches maximum development at ~58 °S, where it is ~500 km across and rises to ~700 m depth Banzare Bank forms the southwest flank of the arch at these latitudes The axis of the arch is marked by a graben which trends almost north–south, the 77° Graben[1], half-grabens on the flanks face towards the central axis The Raggatt Basin (Fig 3), and seismic horizons within it, extend over hundreds of kilometres, and sediment thickness in places is >2 s two-way time[6,7] (twt)

The two major morphotectonic components of the eastern sector are a broad, deep morphological basin (Labuan) and a bounding ridge (William's) These features are characterized by prominent gravity gradients and, in the case of the Labuan Basin, by the broadest negative gravity anomalies of the entire Kerguelen Plateau province Figure 4*a* illustrates the apparent continuity of structure from the horsts and grabens of the north-east margin of the northern and southern sectors into the basins and ridges of the eastern sector Figure 4*b* displays active boundary faults between the eastern and southern plateau sectors The faults coincide with a sharp gravity, gradient, and earthquakes have occurred along trend <35 km north of the seismic line[12,13] The deep basement of the Labuan Basin extends to the southeasternmost extent of the seismic data (~62 °S), and its associated negative gravity signature to the limits of the Seasat-derived gravity data at 63 °S The Labuan Basin seems to extend farther to the Antarctic margin Figure 4*c* shows basement in the basin, it is faulted, averages 7–8 s (twt) in depth, reaches maximum depths of ≥8 s (twt), and has a sediment cover of 0- ≥2 s (twt) thickness Discrete sub-circular strong positive gravity anomalies at 55 5 °S and 57 5 °S are probably seamounts William's Ridge is ~500 km long Profile 2 in Fig 4*a* is the only complete seismic transect of the feature The ridge stands ~1,500 m above the adjacent ocean floor, is ~100 km across, and appears to consist of a series of horst blocks North-east of Labuan Basin and William's Ridge is a sharp gravity gradient which corresponds to changes in basement depth, morphology, and structure[1,6,7] North of ~55 °S the transition observed on seismic data is abrupt[1], and to the south more gradual[6,7] This basement offset (Fig 3) marks the boundary of the eastern sector of the Kerguelen Plateau province with the Tertiary oceanic crust dated by marine magnetic anomalies[2]

The recognition of an eastern sector of the Kerguelen Plateau province and marked differences in the structure of the three plateau sectors implies a more complex evolution of the feature than has been proposed[1,14-16] Specifically, there is a need to review previous fits of the Kerguelen Plateau province and Broken Ridge Previous reconstructions[1,14,15] have employed parameters derived from the separation of Australia and

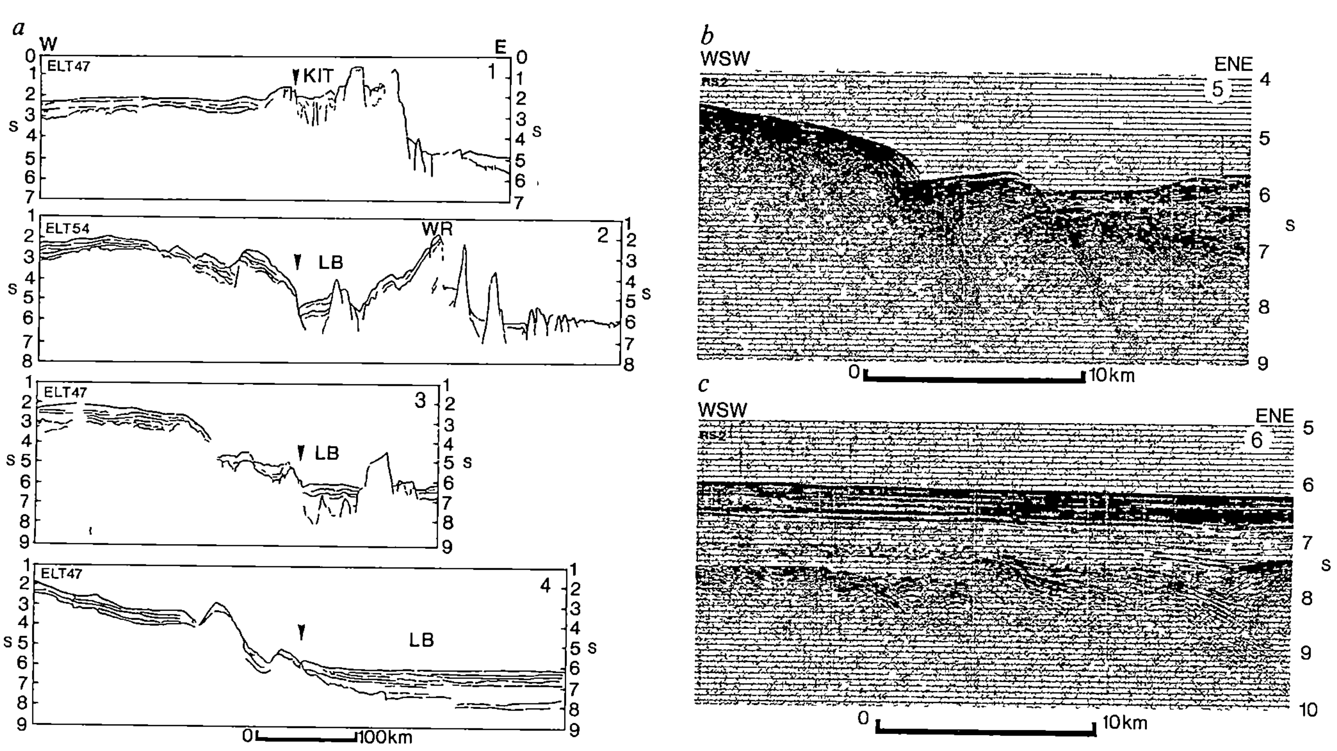


Fig 4 Seismic data documenting the eastern sector of the Kerguelen Plateau province ELT, line drawings of *Eltanin* single-channel seismic data, RS, *Rig Seismic* MCS data Profiles are numbered in upper right hand corner, see Fig 1 *a*, Sequence of line drawings across the north-east margin of the Kerguelen Plateau province Grabens of the Kerguelen Island Trough (KIT)[1] form an apparent rift zone high on the plateau (1), descend (2) into the Labuan Basin (LB, 3), and persist far to the south (4) William's Ridge (WR), previously mapped as an isolated seamount on pre-Seasat bathymetric maps[1,2], is on the north-east end of profile 2, which is the only existing complete transect of the feature The drawings are aligned (black arrows) along the western flank of the Kerguelen Island Trough and Labuan Basin *b*, Seismic line showing detail of major faulting marking the boundary between the southern and eastern plateau sectors Earthquakes in 1973 and 1981 occurred along trend <35 km to the north[13] *c*, Seismic line showing basement character of Labuan Basin

Antarctica, we suggest a new reconstruction employing Seasat-derived fracture zone azimuths between the Kerguelen Plateau and Broken Ridge Dating of the eastern sector and of the Diamantina Zone, presumed conjugates[7], is vital to any tectonic reconstruction for the region

The relationship between the Kerguelen Plateau and hotspot volcanism needs further attention Flexure studies and/or sampling are needed to test if certain apparently compensated components of Kerguelen Plateau province—Elan Bank, Skiff Bank and Kohler seamount—are volcanic edifices whose origins may be unrelated to those of the generally uncompensated main northern, southern, and eastern plateau sectors The rectilinear chain of seamounts between Kerguelen and Heard islands suggests a slow north-west movement of the Antarctic plate in the interval 39–10 Myr, contradicting previously determined Antarctic plate motions[1,17,18] Resolution of the conflicting evidence on absolute motion of the Antarctic plate awaits further sampling and age dating of the volcanic rocks on the plateau The southern plateau sector, in having far greater relief, active major faulting, and no evidence of Neogene volcanism, differs from typical hotspot swells[19] Gaussberg volcano on Antarctica, although chemically distinct from Kerguelen Plateau volcanics[20], lies along the trend of the plateau—is it part of the plateau province? Although further geophysical studies may provide clues, Ocean Drilling Program operations on the Kerguelen Plateau planned for 1987-88 have the most potential for resolving questions of the feature's origin and evolution

We thank the master, crew and the scientific staff of the R/V *Rig Seismic* for their roles in acquiring the multichannel seismic data C Brenner, M Weigel and the BMR Drawing Office provided technical assistance We thank S Cande, J Colwell, P Hill, M Munschy, C Pigram, D Ramsay, R Schlich, J Sclater, H Stagg and B Willcox, for critically reading the manuscript M F C and H L D publish with the permission of the Director, Bureau of Mineral Resources, Lamont-Doherty contribution 4057

Received 2 January, accepted 28 August 1986

1 Houtz, R. E, Hayes, D E. & Markl, R. G *Mar Geol.* **25**, 95-130 (1977)
2 Schlich, R in *The Ocean Basins and Margins 6 the Indian Ocean* 51-148 (Plenum, New York, 1982)
3 Tapley, B D, Born, G H & Parke, M E *J geophys Res* **87**, 3179-3188 (1982)
4 Haxby, W F, Karner, G D, LaBrecque, J L & Weissel, J K *Eos* **64**, 995-1004 (1983)
5 Haxby, W F & Weissel, J K. *J geophys Res* **91**, 3507-3520 (1986)
6 Ramsay, D C *et al Bur Miner Resour Aust Rep* **270**, (1986)
7 Ramsay, D C *et al Geology* **14**, 589-593 (1986)
8 Upton, B J G in *The Ocean Basins and Margins 6 the Indian Ocean* 585-648 (Plenum, New York, 1982)
9 Wicquart, E thesis, *Univ Pierre et Marie Curie, Paris* (1983)
10 Schlich, R, Delteil, J, Moulin, J, Patriat, P & Guillaume, R *C r hebd Seanc. Acad Sci., Paris* **272**, 2060-2063 (1971)
11 Frohlich F *et al C r hebd Seanc. Acad Sci., Paris* **297**, 153-156 (1983)
12 Okal, E A *Earth planet Sci. Lett* **52**, 397-409 (1981)
13 Adams, R D & Zhang, B M *Geophys J R. astr Soc.* **79**, 697-703 (1984)
14 Mutter, J C & Cande, S C *Earth planet Sci. Lett* **65**, 369-376 (1983)
15 Mutter, J C, Hegarty, K. A, Cande, S C & Weissel, J K *Tectonophysics* **114**, 255-279 (1985)
16 Mammerickx, J & Sandwell, D *J geophys Res* **91**, 1975-1988 (1986)
17 Duncan, R A *Tectonophysics* **74**, 29-42 (1981)
18 Goslin, J & Patriat, P *Tectonophysics* **101**, 221-244 (1984)
19 Crough, S T *A. Rev Earth planet Sci.* **11**, 165-193 (1983)
20 Tingey, R J, McDougall, I & Gleadow, A J W *J geol Soc. Austr* **30**, 241-246 (1983)
21 Hayes, D E & Vogel, M *General Bathymetric Chart of the Oceans 5 13 (GEBCO) scale 1 10,000,000* (Canadian Hydrographic Service, Ottawa, 1981)
22 Fisher, R L, Jantsch, M Z & Comer, R L *General Bathymetric Chart of the Oceans 5 09 (GEBCO) scale 1 10,000,000* (Canadian Hydrographic Service, Ottawa, 1982)

Oxygen isotopes and sea level

J. Chappell* & N. J. Shackleton†

* Biogeography and Geomorphology, Research School of Pacific Studies, Australian National University, Canberra, Australia
† Godwin Laboratory for Quaternary Research, University of Cambridge, Free School Lane, Cambridge CB2 3RS, UK

From the time that detailed oxygen isotope records derived from foraminifera living in the constant-temperature environment of the abyssal ocean became available, there has been a discrepancy between the ice volume record that these records imply, and that derived from the altitude of dated coral terraces around the world. Here, we re-examine the data and conclude that the temperature of the abyssal ocean has been an actively varying component of the climate system.

Oxygen isotope ratios in foraminifera from deep-sea sediments have been used as indicators of past climate following Emiliani's pioneering work[1] $\delta^{18}O$ in foraminifera varies with $\delta^{18}O$ in the water from which their carbonate tests are deposited, but differs from the water value by an amount that is temperature dependent[2] The mean $\delta^{18}O$ of the whole ocean varies with the quantity of isotopically light ice stored on the continents so that the record from foraminifera in a particular core is a blend of global (ice volume) and local (temperature) components As the record in planktonic foraminifera may be affected by differences in surface salinity[3] as well as by changing seasonal growth and depth habitat[4], benthic foraminifera have recently been preferred as they provide a less complicated record of global ice volume It has been assumed that the deep ocean, at least in the Pacific, is so cold that its temperature may be regarded as constant as well as spatially uniform

Differences exist between $\delta^{18}O$ records from different cores, even those based on benthic foraminifera Recent studies have isolated systematic differences between good quality benthic records from different regions which have been interpreted as indicating that the deep Atlantic, which at present is ~2 °C warmer than the Pacific, was significantly colder than this during glacial times than at present[5] However, in these studies it was only possible to study inter-oceanic gradients, a change in the average temperature of the whole deep ocean could not be distinguished from a whole-ocean isotopic change

We have made a detailed comparison between a deep Pacific ^{18}O record from core V19-30 (ref 6), and a high-resolution sea-level record from the Huon Peninsula, New Guinea Although neither can be regarded as a perfect monitor of global ice volume, both records must be dominated by a common ice-volume signal In making this comparison, we extend earlier discussions which noted a discrepancy between sea-level records of the past 120 kyr as indicated by raised marine terraces, and by the marine oxygen isotope record[7,8] The record from core V19-30 has been fully documented as has the methodology on which it is based[6] The Huon sea-level record[9] is less well known and its derivation is reviewed to clarify the assumptions involved in comparing it with the ^{18}O record

Sea-level changes over the past 300 kyr have been interpreted from a flight of raised coral terraces along the tectonically rising north-east coast of Huon Peninsula Individual reef terraces were dated using the ^{14}C and $^{230}Th/^{234}U$ methods, which have given good internal consistency within morphological units[10-14] Each reef developed when the rising sea level overtook the rising land, reef crests represent approximately the peaks of each transgression[11,13] Terrace heights increase southeastwards along the coast with increasing uplift rates The morphology and internal structure of each reef indicates the course of sea-level change relative to the rising land, so that sea level relative to stable ocean floor can be extracted for each section if the uplift rate for that section is known

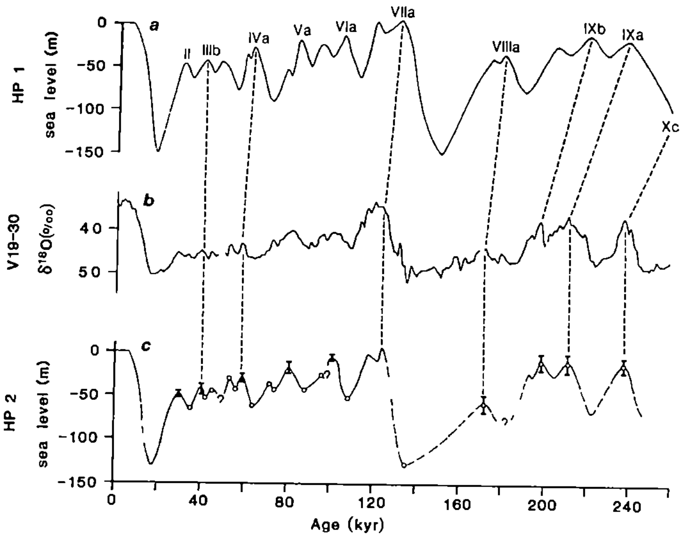


Fig 1 *a*, Detailed sea-level curve for Huon Peninsula from ref 9, with small notches eliminated *b*, ^{18}O record for the past 340 kyr from east equatorial core V19-30, from measurements of benthic *Uvigerina senticosa*, plotted to an orbitally-tuned timescale and smoothed with a three-point running mean *c*, Detailed sea-level curve for Huon Peninsula as recalculated after detailed correlation with the ^{18}O record from core V19-30

We have estimated uplift rates assuming that sea level at ~125 kyr BP was 6 m above the present level[9-12] The spread in radiometric analytical data for this high sea-level event has been interpreted as indicating that this episode may have been prolonged However, stable isotope measurements in both molluscs[15] and corals[16] have confirmed that this interval of high sea level coincided with the rather short substage 5e in the marine oxygen isotope sequence Reef complex VII represents the '125-kyr marker' in the Huon sequence, and is divided into units VIIa and VIIb[9-12,17]

Sea levels associated with other Huon reefs were estimated as follows given the uplift rate U_j based on the elevation of reef VII in section j, sea level $S_{i,j} = H_{i,j} - U_j t_i$ where $H_{i,j}$ is the elevation of the crest of reef i on section j and t_i is the age of reef i Sea levels have been estimated for the reef sequence at seven surveyed sections with uplift rates ranging from 0 9 m kyr^{-1} (north-west end of the flight) to 3 5 m kyr^{-1} (south-east end of the flight)[12] Error estimates for the sea level associated with each reef crest are based on results from all sections, combined with the contribution of dating errors[9] The lower sea levels which intervene between reef crests were estimated in the same way, using the heights of shallow marine and littoral deposits in the Tewai section[9], supported by dates from other sections[13,14]

The detailed sea-level curve for the past 260 kyr derived by this method is shown in Fig 1*a*, modified from the previous version only in that small steps of a few metres amplitude have been smoothed out These steps were interpreted from wave-cut notches throughout the terrace flight, they seem to represent individual seismic events and thus have no global validity[9] The case for regarding Fig 1*a* as a global curve rests on points of agreement with results from elsewhere Sea levels at 105, 83 and 60 kyr agree closely with estimates from Barbados[12] and Timor[13] as well as being supported by dates from the Ryukus[19] and in Vanuatu (New Hebrides)[20] Sea level was estimated in these other areas using the same method, with the reef or terrace of 125 kyr being the reference surface in each case Agreement between sites which are widely separated and which have very different uplift rates supports the global use of the Huon curve However, less support is available for the low points, because

Table 1 Heights, uplift rates and sea-level estimates for Huon Terrace sections

Huon Reef	Age (kyr adjusted)	Tewai	Kanzarua	Blucher	Kwambu	Nama	Sambero	Mean sea level
VIIa	124	440	330	280	220	160	150	+6 (assumed)
Uplift (m kyr^{-1})		3.5	2.61	2.21	1.72	1.24	1.16	(from VIII a)
VIa	100	338*	250	215	160	115	110	
		−12	−11	−6	−12	−9	−6	−9±3
VIb	96	312						
		−24						
Va	81	260	190	155	117	90	80	
		−23	−21	−24	−22	−10	−14	−19±5
Vb	72	216						
		−36						
IVa	59	178	125		70	48		
		−29	−29		−31	−25		−28±3
IVb†	53	156						
		−30						
IIIa	45	112						
		−46						
IIIb	40	98	70	41	28	10		−41±4
		−42	−34	−47	−41	−40		
II	28	52	30	18	7			
		−46	−43	−44	−41			−44±2

Height data are from ref. 9 (Tewai) and ref. 12 (other sections). Heights of reef VIIa at all except Tewai and Kwambu are as given for VIIb in ref. 12. Uplift rate is based on VIIa = 12 kyr, sea level at +6 m.

* Estimated, not surveyed, height. All other heights are from theodolite survey.

† Heights of IVb and IIIa based only on Tewai section, as uncertainty exists over whether IIIa surveyed on other sections is equivalent to 53- or 45-kyr isotope points.

the low sea-level deposits are generally buried by those of the subsequent transgression. Low sea-level points are a good test of the global applicability of the curve because isostatic and other factors will probably have their greatest effect between glacial maxima and minima.

The low sea level at 18 kyr corresponding to the last glacial maximum is an important point for comparison with ^{18}O records. Figure 1*a* shows this level at −150 m based mainly on results from northern Australia where lagoonal facies at −135 m (ref. 21), intertidal beach rock at −150 m (ref. 22), and coral below a terrace at −165 m (ref. 23) occur at sites on the outer shelf margins and are dated between 14 and 18 kyr. Lesser values, around −90 to −110 m are reported from the west Atlantic[24] and Texas Gulf[25]. Variations around the globe are to be expected as a result of adjustment to glacial/interglacial changes of ice and ocean volumes[26-28], although the magnitude of predicted variation differs from one geophysical model to another. The sea-level change between 18 kyr and the present on a broad shelf would be expected to be larger than that recorded on a dipstick-type margin such as the Huon Peninsula[28-30], so that 150 m may be an overestimate of the 18-kyr level at Huon. Undated evidence does occur at Huon in the form of a submerged terrace running along the submarine margin. This is narrow and coraline where it occurs in front of coral terraces, and broad and gravelly where it lies off the flank of a broad deltaic plain[11]. Correcting its level for local uplift puts the sea level at −130 m for each locality[11]. In view of this consistency, we use this estimate of −130 m in our subsequent discussion.

The $\delta^{18}O$ record from core V19-30, based on *Uvigerina senticosa*, is shown in Fig. 1*b*. The original sampling interval of 3 cm corresponds to ~500 yr on average although this value varies down the core as accumulation rate varies[6,31]. Figure 1*b* shows the record using a timescale that was developed by the SPECMAP project[32], and smoothed using a three-point running mean. This timescale was developed iteratively, the starting point being radiometric age determinations both in marine sediments and in coral terraces. The final timescale is developed by tuning the initial record on the basis of its relationship to orbital precession, obliquity and eccentricity functions[32,33]. Although the final timescale is not dependent in detail on particular radiometric dates, the modifications are within the uncertainties of the age estimates at all points[32,33]. In fact, the orbital theory does not provide sufficient constraints on the timescale within stage 3 (in the range 30-60 kyr BP) so that no significant departures from the chronology implied by direct radiometric dating have been used in this interval.

We consider that for the part of the record older than stage 3, the orbitally tuned timescale should be used to refine the age estimates for the Huon Peninsula reefs, on the grounds that independent lines of evidence converge in support of this mechanism, and that the assumption of orbital forcing places constraints on the timescale from 70 kyr back past the useful limits of U-series dating that are much tighter than those imposed by the analytical uncertainties on radiometric dates in this interval. Figure 1*c* shows the Huon sea-level record as re-estimated on this basis, each sea-level maximum and minimum has been assigned the same age as the assumed correlative in the V19-30 record. The amplitude of some peaks differs significantly between Fig. 1*a* and *c* because the sea-level estimate for any point in the sequence depends on the assumed age for that point. The sea level also depends on the value of the uplift rate adopted in each section which depends on the age assumed for reef VII. It also depends on whether this age, and the 6-m calibration height, is assigned to terrace VIIa or VIIb. Finally in traverses containing a single feature for reef VII the uplift rate depends on whether this feature is assumed to represent VIIa or VIIb. Because oxygen isotope studies clearly associate reef VIIa (the older) with substage 5e (ref. 17), and also associate the undifferentiated but dated level in Barbados[15,16] and in Curacao with 5e, it must be assumed that VIIa is the calibration level to make a comparison consistent with the deep-sea ^{18}O record. There are various sources of data from the Huon Peninsula[9-14], so to remove confusion, Table 1 lists reef crest heights from each surveyed traverse together with the age that we have adopted for each, and the resulting sea-level estimate.

Before 130 kyr the sea-level curve derived from marine terraces is subject to larger uncertainties because both the assigned ages and assumed uplift rates become less secure. However, for

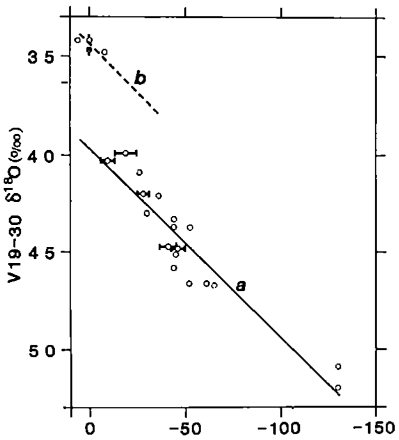


Fig 2 Huon sea level (recalculated age model, Fig 1*c*) versus equivalent ^{18}O in V19-30, from data in Table 2 Regression line *a*, $\delta^{18}O = 3\ 96 - 0\ 0097$ sea level, fitted to all points from 17 to 112 kyr, $r^2 = 0\ 86$ Line *b* is drawn parallel to the regression line *a*, but passes the area at top left which encompasses data for 5e (124 kyr) and recent (6 kyr-present)

the past 130 kyr we regard Fig 1*c* as the best sea-level record available for the New Guinea/Australian region (noting that the magnitude of sea-level oscillations may be smaller in other parts of the world as a result of isostatic factors) Within the past 130 kyr the greatest uncertainty relates to the reef IVb-IIIa area where the isotopic record shows little structure (see Table 2), in Fig 1*c* a gap is shown over this part Table 2 lists sea levels for most turning points in Fig 1*a*,*c* as well as the ^{18}O values for corresponding points in V19-30

Comparing Fig 1*b* with Fig 1*c*, it is particularly striking that the sea-level peaks within the last glacial cycle, and especially the well controlled peaks at ~106 and 81 kyr, are relatively closer to the Recent and 124-kyr levels than the equivalent ^{18}O values Figure 2 plots sea level against ^{18}O for all established turning points in Fig 1*c* A cluster of points around zero sea level and +3 4‰ corresponds to full interglacial stage 1 and substage 5e A single line cannot be drawn through all data points that is, there is not a simple linear relationship between sea level and ^{18}O in benthic foraminifera Although this situation could result in part from temporal variations in the isotopic composition of the stored ice, this cannot be the chief factor involved If the 0 6‰ isotopic difference between the interglacial points and the two terraces at ~−20 m were to be explained solely in terms of the removal of isotopically light water from the ocean, this would necessitate ice with an ^{18}O content of ~100‰ The lightest ice in Antarctica is ~−60‰[34,35] The average isotopic composition of the ice in an ice sheet depends mainly on the isotopic composition of the snow falling in the cold central region, which depends on the air temperature and hence mainly on the altitude of the central part Thus the first ice accumulating in small Northern hemisphere ice sheets must have been less isotopically light than was the ice stored in the fully developed ice sheet

One possible solution to this difficulty could be that a significant amount of isotopically light floating ice accumulated[36-38] This would change the ocean isotopic composition without having any effect on sea level The only place where a sufficient quantity of floating ice can plausibly be envisaged is the Arctic Ocean, the detailed and continuous stratigraphic sequences that have recently been documented from deep-sea cores in the Eastern Arctic Basin[39] exclude the possibility that this ingenious idea is correct, at least for the 18 kyr glacial maximum The data in Fig 2 require a mechanism that operated during warm substages 5a and 5c as well as at glacial maximum time

A more probable explanation is that this isotopic shift results from a temperature effect on the isotopic composition of the benthic foraminifera analysed A best-fit line through all the glacial points in Fig 2 intercepts the axis at ~+3 95‰ (line *a* on Fig 2), requiring a temperature difference of ~2 °C The slope of this line is 0 97‰ per 10 m (very close to the generally adopted value[40]) which corresponds to a mean ^{18}O content for the abstracted water of ~−35‰, a reasonable value

Deviations from this regression are still significant, in Fig 3 we examine the temporal pattern of the deviations from line *b* on Fig 2, which parallels line *a* and passes the Recent points, to evaluate the relative contributions of varying sea level, deep-water temperature and ice isotopic composition to the ^{18}O record For convenience, Fig 3 has a scale representing deep-water potential temperature at the site of core V19-30 (refs 41, 42) This scale would be valid only if there were no variation in the isotopic composition of stored ice-sheet ice If this assumption is correct, Pacific deep waters cooled rapidly to ~0 °C by ~110 kyr, and then cooled more slowly to ~−1 °C around 40 kyr, warming to 0 °C at 18 kyr and then to 1 5 °C at present Although a warming at 18 kyr seems surprising, one model has predicted such an effect[43] However, an alternative would be that the stored ice (which includes Antarctic ice) became gradually more negative between 110 and 30 kyr (ref 44), and that the very rapid advance of ice to lower latitudes between 30 and 20 kyr involved more temperate ice that was less isotopically negative Further work will be needed to establish the possible contribution of temperature variability to the pattern in the glacial section of Fig 3, we regard the overall glacial cooling of Pacific deep water by 1 5 °C that is implied, as inescapable

It would be useful to compare the older sections of the Huon and V19-30 records, but this is not possible for several reasons (1) there is no earlier sea-level datum comparable to that at 124 kyr to use for uplift-rate calibration, (2) the errors in $^{230}Th/^{234}U$ dates become more serious in this interval, Fig 1 suggests that this method may actually overestimate ages in the 180-240 kyr age range, (3) only two sections have been surveyed for the pre-5e part of the Huon record It is regrettable that the morphostratigraphy of this important Upper Quaternary area has been surveyed and described by only one geologist (J C)

Table 2 Huon sea levels and V19-30 ^{18}O values

Huon Reef	Age HP1 (kyr)	Sea level HP1 (m)	Age HP2 (kyr)	Sea level HP2 (m)	$\delta^{18}O$
Modern	0	0	0	0	3 42
I	6	0	6	0	3 47
I/II	18	−150	17	−130	5 09
II	30	−46	29	−46	4 48
II/IIIb	34	−62	35	−65	4 67
IIIb	40	−40	40	−41	4 47
IIIb/a	43	−57	42	−52	4 66
IIIa	47	−42	44	−45	4 51
IIIa/IVb	55	−78	?	?	—
IVb	58	−36	53	−30	4 30
IVb/a	60	−40	56	−44	4 58
IVa	62	−27	59	−28	4 20
IVa/Vb	72	−88	64	−61	4 66
Vb	77	−52	72	−36	4 21
Vb/a	79	−60	74	−44	4 33
Va	83	−18	81	−19	3 99
Va/VIb	89	−48	88	−44	4 37
VIb	95	−22	96	−26	4 09
VIb/a	99	−38			
VIa	105	−12	106	−19	4 03
VIa/VIIb	112	−60	112	−62	4 38
VIIb	120	+4	118	0	3 49
VIIb/a	123	−8	122	−8	—
VIIa	132	+6	124	+6	3 37
VII/VIII	150	−145	135	−130	5 20

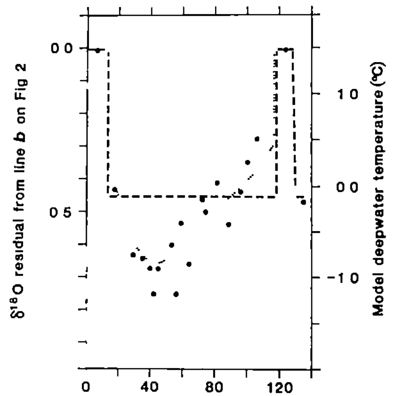


Fig 3 Residuals from line *b* (Fig 2) versus estimated age Scale to right shows calculated potential temperature at the sea floor at the site of core V19-30, if this line is interpreted as the correct relationship between sea level and ocean $\delta^{18}O$, and deviations from the line ascribed to temperature Heavy dashed line is our preferred interpretation of the temperature record, the light dotted line is an alternative

No other area is known where the details of sea-level change are available for field examination

We conclude that deep waters in the Pacific Ocean were ~1 5 °C cooler in glacial and interstadial times than in the short (~10 kyr duration) interglacials of substage 5e and the present One interpretation of our analysis is that deep water was coldest at ~40 kyr and that almost 1 °C of warming occurred by the glacial maximum and a further 1 5 °C of warming subsequently Alternatively, part of the observed detail resulted from changes in the isotopic composition of the ice sheets, with the deep Pacific remaining ~1 5 °C cooler than the present temperature throughout the glacial Regardless of this detail, massive cooling of the already cold abyssal ocean must be considered an important aspect of the ice age climate

This work was supported in Cambridge by the NERC N J S acknowledges stimulating discussion with many colleagues, particularly J -C Duplessy

Received 26 March, accepted 22 August 1986

1 Emiliani, C *J Geol* **63,** 538-538 (1955)
2 Epstein S , Buchsbaum, R., Lowenstam, H A & Urey, H C *Bull geol Soc Am* **62,** 417 (1951)
3 Epstein S & Mayeda, T *Geochim cosmochim Acta* **4,** 213-224 (1953)
4 Berger, W H , Killingley, J S & Vincent, E *Oceanol Acta* **1,** 221-230 (1978)
5 Duplessy, J C , Moyes, J & Pujol, C *Nature* **286,** 479-481 (1980)
6 Shackleton, N J , Imbrie, J & Hall, M A *Earth planet Sci Lett* **65,** 233-244 (1983)
7 Broecker, W S *Science* **188,** 1116-1121 (1975)
8 Dodge, R E , Fairbanks, R G , Benniger, L K & Maurrasse, F *Science* **219,** 1423-1424 (1983)
9 Chappell, J *Search* **14,** 99-101 (1983)
10 Veeh, H H & Chappell, J *Science* **167,** 862-865 (1970)
11 Chappell, J *Bull geol Soc Am* **85,** 553-570 (1974)
12 Bloom, A L., Broecker, W S , Chappell, J , Matthews, R K & Mesolella, K J *Quat Res* **4,** 185-205 (1974)
13 Chappell, J & Polach, H A *Bull geol Soc Am* **87,** 235-240 (1976)
14 Chappell, J & Veeh, H H *Nature* **276,** 602-603 (1978)
15 Shackleton, N J & Matthews, R. K *Nature* **268,** 618-620 (1977)
16 Fairbanks, R. G & Matthews, R K *Quat Res* **10,** 181-196 (1978)
17 Aharon, P , Chappell, J & Compston, W *Nature* **283,** 649-651 (1980)
18 Chappell, J & Veeh, H H *Bull geol Soc Am* **89,** 356-368 (1978)
19 Kohishi, K , Omura, A & Nakamichi, O *Proc 2nd int Coral Reef Symp* **2,** 595-613 (1973)
20 Neef, G & Veeh, H H *Nature* **269,** 682-683 (1977)
21 Van Andel, Tj & Veevers, J J *Bur Miner Res , Aust Bull* **3** (1967)
22 Jongsma, D *Nature* **228,** 150-151 (1970)
23 Veeh, H H & Veevers, J J *Nature* **226,** 536-537 (1978)
24 Milliman, J D & Emery, K. O *Science* **162,** 1121-1122 (1970)
25 Curray J *Recent Sediments, Northwest Gulf of Mexico* (ed Sheppard, F P) 221-266 (American Association of Petroleum Geologists, Tulsa, 1960)
26 Walcott, R I , *Quat Res* **2** 1-14 (1970)
27 Cathles, L M in *Earth Rheology, Isostasy and Eustasy* (ed Morner, N A) 11-44 (Wiley, New York, 1980)
28 Clark, J A , Farrell, W E & Peltier, W R *Quat Res* **9,** 265-287 (1978)
29 Bloom, A L *Bull geol Soc Am* **78,** 1477-1494 (1967)
30 Chappell, J *Quat Res* **4,** 405-428 (1974)
31 Shackleton, N J & Pisias, N G in *Atmospheric CO_2 Natural Variations, Archaean to Present* (eds Sundquist, E T & Broecker, W S) 303-317, Am geophys Un monogr **32** (1985)
32 Imbrie, J *et al* in *Milankovich and Climate* (eds Berger, A *et al.*) 269-305, (Reidel, Dordrecht 1984)
33 Hays, J D , Imbrie, J & Shackleton, N J *Science* **194,** 1121-1132 (1976)
34 Dansgaard, W , Johnson, S J , Moller, J & Langway, C C *Science* **166,** 377-381 (1969)
35 Picciotto, E , Cameron, R , Crozaz, G , Deutsch, S & Wilgain, S *J Glaciol* **7,** 273-287 (1968)
36 Mercer, J H *Palaeogeogr , Palaeoclimatol Palaeoecol* **8,** 19-27 (1970)
37 Broecker, W S *Science* **188,** 1116 (1975)
38 Williams, D F , Moore, W S & Fillon, R H *Earth planet Sci Lett* **56,** 157-166 (1981)
39 Markussen, B , Zahn, R & Thiede, J *Palaeogeogr Palaeoclimatol., Palaeoecol* **50,** 271-284 (1985)
40 Shackleton, N J & Opdyke, N D *Quat Res* **3,** 39-55 (1973)
41 O Neil, J R , Clayton, R N & Mayeda, T K. *J chem Phys* **51,** 5547-5558 (1969)
42 Shackleton, N J *Colloq Int cent Nat Res Scient* **219** (1974)
43 Newell, R *Quat Res* **4,** 117-126 (1974)
44 Mix, A C & Ruddiman, W F *Quat Res* **21,** 1-20 (1984)

Predicting aqueous aluminium concentrations in natural waters

Christopher S. Cronan, William J. Walker* & Paul R. Bloom†

Department of Botany and Plant Pathology, University of Maine, Orono, Maine 04469, USA
* Land and Water Resources Center, University of Maine, Orono, Maine 04469, USA
† Department of Soil Science, University of Minnesota, St Paul, Minnesota 55108, USA

Aluminium is a *p*H-sensitive element that can cause acute toxicity symptoms in some organisms at aqueous activities of 10 μM or less[1–3]. Scientists working on agricultural systems have long been concerned with the deleterious effects of aluminium on crop roots[4,5]. More recently, environmental scientists have reported a potentially harmful biogeochemical link between acidic deposition onto forest soils and aluminium toxicity in forest and aquatic communities of northeastern North America and northern Europe[6–8] Because of this general interest in aluminium toxicity as an environmental threat, there have been renewed efforts to model the chemistry and transport of aqueous aluminium in soils and surface waters. Here we propose that much of the spatial and temporal variability in aqueous aluminium chemistry can be accounted for by a two-component equilibrium model involving a solid-phase humic adsorbent and an aluminium trihydroxide mineral phase. Inputs for the model are solution *p*H, copper-extractable organic aluminium and the titratable carboxyl content of soil humus.

Several workers[9,10] have found that [Al^{3+}] (aquo aluminium ion activity) in surface waters can be predicted accurately as a function of *p*H using an aluminium trihydroxide solubility relationship However, other reports[11–13] have presented data revealing aluminium concentrations that are much lower than predicted—that is, despite *p*H values less than 4 5 and elevated concentrations of complexing organic ligands, the solutions contained low concentrations of soluble aluminium This illustrates the failure of existing geochemical models to account for all the major variations in [Al^{3+}] and we refer to it as the 'paradox of aluminium undersaturation'

Figure 1 shows [Al^{3+}] for a variety of natural waters from North America and northern Europe The expected pAl^{3+} values for solutions in equilibrium with natural gibbsite and amorphous $Al(OH)_3$ are also shown As illustrated, there are three major points where observed and predicted aqueous aluminium activities diverge the upper soil rooting zone, downstream of wetlands and during stream peak flows While it has been suggested that these cases of lowered [Al^{3+}] are the result of kinetic constraints on the dissolution of mineral aluminium[14], there have been no rigorous efforts to confirm the kinetic hypothesis in natural systems In this study, our objective was to derive a model that could account for these 'undersaturated'

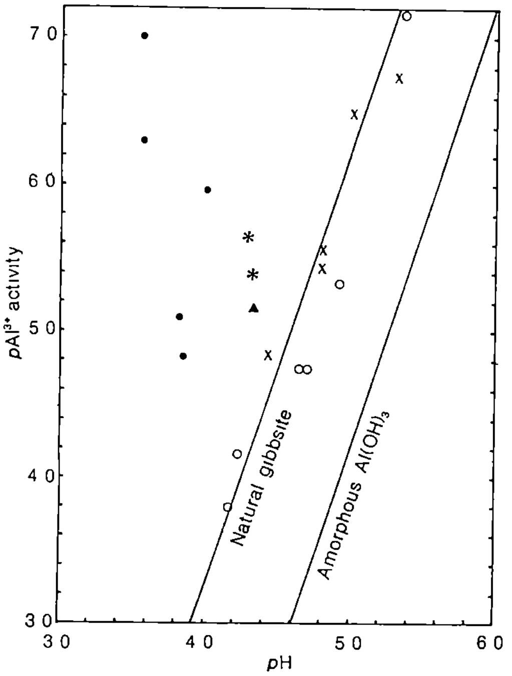


Fig 1 Aquo aluminium activity as a function of pH for a variety of natural waters from North America and Europe Field samples include upper O/A horizon soil solutions (●), lower B/C horizon soil solutions (○), wetland drainage (∗), stream storm flow (▲), and stream baseflow (×) from the following watersheds Woods Lake, New York, Huntington Forest, New York, Pancake-Hall Creek, New York, Solling spruce forest, West Germany, Plastic Lake, Ontario, Lange Bramke, West Germany, Camp Branch, Tennessee, Birkenes, Norway and Lake Gardsjon, Sweden (C S C and C L Schofield, unpublished data)

waters, so that [Al^{3+}] could be predicted across the range of inorganic and organic environments illustrated in Fig 1

Soil and solution chemistry data for this study were obtained using methods described in Table 1 Laboratory experiments were conducted with field-moist, 2 mm sieved forest floor horizons (Oi and Oa)[15] from contrasting northern and southern watersheds in the United States and Canada Aqueous aluminium ion activities were estimated using the GEOCHEM program[16], appropriate thermodynamic constants[17], and the following mass balance equation

$$Al^{3+} = \mathrm{MAL} - \mathrm{MOAL} - Al(OH)_x^{3-x^+}$$
$$-(\text{complexes of } F^- \text{ and } SO_4^{2-})$$

where,

MAL = total monomeric aluminium

MOAL = organic monomeric aluminium

As shown in Fig 1, many of the data points from natural waters lie between the solubility lines for gibbsite and amorphous $Al(OH)_3$ This general pattern agrees with the results of previous studies[10,18] and implies that [Al^{3+}] in these waters may be largely controlled by an aluminium trihydroxide solid phase In order to model solution aluminium chemistry for these types of waters, one can take an average solubility product of $\log K^0 = 8\ 77$, corresponding to natural gibbsite[18] or interlayer aluminium hydroxide[19] Then [Al^{3+}] in solution can be related to the aluminium hydroxide solid-phase by the following relationship

$$\log[Al^{3+}] = 8\ 77 - 3\,pH \qquad (1)$$

Combining this with measured concentrations of complexing ligands and appropriate thermodynamic data[17], it is possible to estimate the equilibrium distribution of inorganic aqueous aluminium

Table 1 Apparent solubility expressions for predicting Al^{3+} activity in the presence of soil organic matter at different bound aluminium ratios

Bound aluminium ratio	Slope	Intercept	Solubility expression	pAl^{3+} at pH 4 40
0–0 10	0 40	3 62	$pAl = 0\ 40(pH) + 3\ 62$	5 38
0 20–0 30	0 60	2 52	$pAl = 0\ 60(pH) + 2\ 52$	5 16
0 60–0 70	1 04	0 26	$pAl = 1\ 04(pH) + 0\ 26$	4 84
0 90–1 00	1 34	−1 28	$pAl = 1\ 34(pH) - 1\ 28$	4 62

Calculated values of pAl^{3+} at pH 4 40 are illustrated for each case Regressions are based on Oi and Oa samples from a representative Adirondack Spodosol Organically bound Al in each soil horizon was determined by extraction with 0 5 M $CuCl_2$, followed by ICP-AES analysis of Al (ref 21) Subsamples of each soil were hydrogen saturated with 1 M HCl[11] and were then titrated under nitrogen to pH 7 0 in 2×10^{-3} M $CaCl_2$ with 0 01 M NaOH Titratable acidity to pH 7 was used as an operational estimate of RCOOH groups The bound Al ratio for each native soil horizon was computed as the equivalents of Cu-extractable Al per unit organic matter divided by the equivalents of RCOOH per unit organic matter Based upon the titration data, subsamples of Oi and Oa forest floor horizons were amended with $Al(NO_3)_3$ to produce a range of bound Al ratios from 0 1 to 1 0 For each soil horizon and bound Al ratio, subsamples were augmented with 2×10^{-3} M $CaCl_2$ and were adjusted with HCl or NaOH to yield a pH range from 3 0 to 5 0 In addition, untreated control subsamples of all native soil horizons were prepared as soil pastes in H_2SO_4(pH 4 0) augmented with 2×10^{-3} M $CaCl_2$ All soil mixtures were then equilibrated for 48 h, followed by mechanical vacuum extraction through acid washed cellulose pulp Aqueous monomeric inorganic and organic Al fractions were estimated using ion exchange separation, reaction with 8-hydroxyquinoline at pH 8 3, and spectrophotometric determination at 395 nm[22] Inorganic anions were measured with ion chromatography

Earlier studies have emphasized the potential importance of soil organic matter in controlling aqueous aluminium, and Bloom *et al*[11] suggested that aluminium-hydrogen ion exchange on organic matter controls [Al^{3+}] in acidic surface soils We propose that a general model for predicting aqueous aluminium chemistry must incorporate this type of organic solid phase reaction to account for the zone of extreme undersaturation illustrated in Fig 1

Soil equilibration experiments were used to determine the influence of humic adsorption-desorption reactions on [Al^{3+}] in Oi and Oa organic soil horizons Results indicated that (1) forest floor saturated pastes reached stable levels of [Al^{3+}] in 5 min or less and (2) at solution pH values between 3 0 and 5 0, plots of pAl^{3+} versus pH were approximately linear for discrete ranges of bound aluminium, and [Al^{3+}] increased with increasing percentages of solid-phase aluminium (Fig 2) This pattern indicates that equilibrium conditions can be assumed for these soil horizons and that it is possible to construct [Al^{3+}] apparent solubility lines based upon the bound aluminium ratio and pH

Table 1 shows the best-fit linear regression solubility expressions for the four experimental ranges of bound aluminium ratio shown in Fig 2 Combining these equations, gives the following general solubility expression for predicting aquo aluminium ion activity in natural waters dominated by humic adsorption-desorption reactions

$$pAl^{3+} = b(pH) + c \qquad (2)$$

where b is the slope = 1 05 × (bound Al ratio) + 0 345, and c is the intercept = −(5 47 × (bound Al ratio)) + 3 879 This empirical regression equation confirms the strong relationship between [Al^{3+}], solution pH and carboxyl-bound Al, and provides a simple means of predicting [Al^{3+}] in acidic organic soil horizons

The same data for aluminium solubility can be modelled more formally using the modified mass action expression derived by Langmuir[20] Given the reaction, $3H^+ + AlX \rightleftharpoons H_3X + Al^{3+}$, the empirical exchange function would be $K_{ex} = ([Al^{3+}]/[H^+]^3)(H_3X/AlX)^n$, where K_{ex} and n are constants, the brackets denote activities of the ions, and AlX and H_3X are their mole fractions on organic sorbent X Log-linearization of this expression yields

$$\log([Al^{3+}]/[H^+]^3) = \log K_{ex} + n \log(AlX/H_3X) \qquad (3)$$

Table 2 Model validation for a variety of North American forest soils

Watershed	Horizon	pH	Cu Al	RCOOH	Bound Al ratio	pAl^{3+} observed	pAl^{3+} calculated RCOO-Al	pAl^{3+} calculated $Al(OH)_3$
			milliequivalents per gram (Ash-free basis)					
Pancake-Hall, New York	Oi	5 00	0 06	1 77	0 04	5 37	5 60*	6 23
	Oa	4 06	0 09	3 13	0 03	5 30	5 25*	3 41
	Bh	4 35	5 52	5 88	0 94	4 30	4 53*	4 28†
	Stream	4 80	—	—	—	5 55	—	5 63†
Huntington, New York	Oi	5 21	0 10	2 10	0 05	5 70	5 70*	6 86
	Oa	4 11	0 80	2 40	0 34	4 51	4 90*	3 56
	Bh	4 41	12 17	12 64	0 96	4 87	4 59*	4 46†
Plastic Lake, Ontario	Oi	5 32	0 03	1 06	0 03	5 81	5 72*	7 19
	Oa	4 55	0 09	2 02	0 04	5 39	5 42*	4 88
Woods Lake, New York	Oi	4 40	0 07	1 75	0 04	5 59	5 37*	4 43
	Oa	4 02	1 11	5 20	0 21	4 86	4 99*	3 29
	Bh	4 09	1 52	1 62	0 93	4 57	4 20*	3 50
Camp Branch, Tennessee	O	5 20	0 17	2 17	0 08	5 99	5 67*	6 83
	Stream	5 02	—	—	—	6 49	—	6 29†
Univ Forest, Missouri	O	5 30	0 04	0 84	0 05	6 54	5 71	7 13

Observed values for pAl^{3+} were obtained with soil paste extracts, while calculated pAl^{+3} values were derived with either the humic-aluminium or aluminium trihydroxide model components The former model (equation 2) was calibrated with experimental data from the Pancake-Hall Adirondack watershed, the apparent solubility lines from that Spodosol were broadly representative of the other soils

* Close agreement (±0 4) between pAl^{3+} observed and humic-Al model

† Close agreement (±0 4) between pAl^{3+} observed and inorganic solubility model

Table 2 contains a summary of the model validation results obtained for a number of soils from North America with bound aluminium ratios ranging from 0 03 to 1 00 Using the actual bound aluminium ratio and solution pH for each soil sample, the humic-aluminium partition model was used to predict solution $[Al^{3+}]$ for the untreated native soil horizons from each watershed These predictions were compared with observed soil paste $[Al^{3+}]$ values and with predictions from the $Al(OH)_3$ solubility model The fit of a model was judged satisfactory if it correctly predicted pAl^{3+} to within ±0 4 units The humic-aluminium model was successful in predicting the pAl^{3+} in all the O horizons with the exception of the sample from the Missouri watershed Whereas the $Al(OH)_3$ solubility model provided an accurate prediction for pAl^{3+} in stream water at low flow, the model was unable to predict pAl^{3+} correctly in any of the O horizons

We examined the region of the stability field where the humic-aluminum and $Al(OH)_3$ apparent solubility lines intersect The Oi horizon data in Table 2 shows that surface organic horizons with higher pH (>4 8 to 5 0) and low mineral content will often follow the humic-aluminum solubility line and can actually maintain oversaturation with respect to the solubility of interlayer $Al(OH)_3$ This is probably due to the scarcity of clay nucleation sites in these horizons and interference of soluble organic matter in the formation of mineral precipitates However, as the pH is raised above about 5 2, it is likely that amorphous $Al(OH)_3$ forms by homogeneous nucleation Thus, the upper limit of applicability for the humic-aluminium model is probably about pH 5 2

In practice, this model will typically apply to surface organic horizons and peaty hydrologic source areas, whereas the inorganic solubility model will be appropriate for lower mineral soil horizons and for the groundwater zone During peak flows in lower order streams, it is likely that stream water chemistry will be strongly influenced by 'undersaturated' lateral flow derived from organic-rich surface soil horizons Under these conditions, where streamwater results from mixing of lateral flow and baseflow, it will be necessary to combine accurate hydrologic flux estimates with these chemical models

The model presented here can account for much of the variation in $[Al^{3+}]$ observed in natural waters draining different inorganic and organic environments Although we previously suspected that waters undersaturated in $[Al^{3+}]$ could result from kinetic constraints on Al dissolution and desorption, our results suggest that these waters may actually be in equilibrium with a solid phase humic adsorbent With further testing and development, this model may be suitable for application to studies of aluminium transport associated with acid deposition

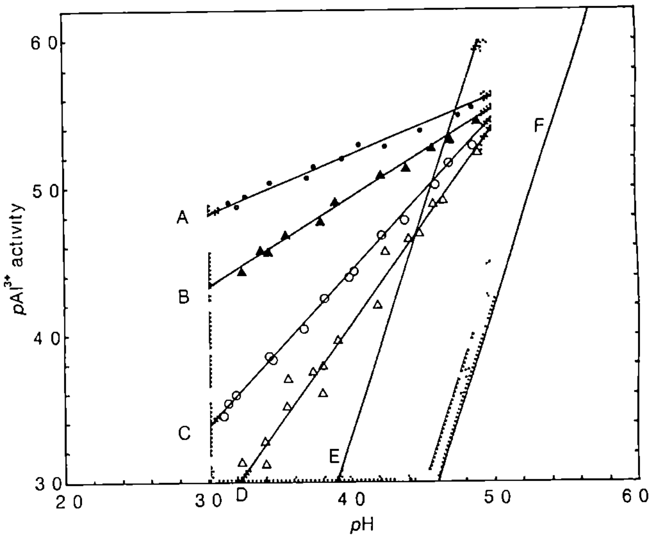


Fig 2 Stability field for apparent aluminium solubility in the presence of soil humic material in the pH range 3 0 to 5 0 The points along each apparent solubility line represent equilibrium pAl^{3+} values for Oi and Oa horizons from an Adirondack Spodosol These horizons were pretreated to yield the following four different ranges of bound aluminium ratio line A = 0 to 0 1, line B = 0 2 to 0 3, line C = 0 6 to 0 7, and line D = 0 9 to 1 0 The solubility lines for interlayer (E) and amorphous $Al(OH)_3$ (F) at 25 °C are shown for reference

We thank C L Schofield for contributing to many of the Al speciation measurements that provoked this investigation and B Lee for typing the manuscript This is a contribution from the ALBIOS project, supported by contract RP 2365 from the Electric Power Research Institute

Received 13 January, accepted 6 August 1986

1 Haug, A *CRC Crit Rev Pl Sci* **1**, 345-373 (1984)
2 Baker, J P & Schofield, C L *Wat Air, Soil Pollut* **18**, 289-309 (1982)

3. Hue, N. V., Craddock, G. R. & Adams, F. *Soil Sci. Soc. Am. J.* **50,** 28–34 (1986).
4. Hardy, F. *J. agric. Sci.* **16,** 616–630 (1926).
5. Pavan, M. A. & Bingham, F. T. *Soil Sci. Soc. Am. J.* **46,** 993–997 (1982).
6. Cronan, C. S. & Schofield, C. L. *Science* **204,** 304–306 (1979).
7. Ulrich, B., Mayer, R. & Khanna, P. K. *Soil Sci.* **130,** 193–199 (1980).
8. Henriksen, A., Skogheim, O. K. & Rosseland, B. O. *Vatten* **2,** 255–260 (1984).
9. Christophersen, N., Seip, H. M. & Wright, R. F. *Wat. Resour. Res.* **18,** 977–996 (1982).
10. Nordstrom, D. K. & Ball, J. W. *Science* **232,** 54–56 (1986).
11. Bloom, P. R., McBride, M. B. & Weaver, R. M. *Soil Sci. Soc. Am. J.* **43,** 488–493 (1979).
12. Cronan, C. S. *Oikos* **34,** 272–281 (1980).
13. Driscoll, C. T., Van Breemen, N. & Mulder, J. *Soil Sci. Soc. Am. J.* **49,** 437–444 (1985).
14. Hooper, R. P. & Shoemaker, C. A. *Science* **229,** 463–465 (1985).
15. Soil Survey Staff *Soil Taxonomy.* (Soil Conservation Service, USDA Washington D.C., 1981).
16. Mattigod, S. V. & Sposito, G. *Am. chem. Soc. Symp. Ser.* **93,** 837–856 (1979).
17. Nordstrom, D. K., Valentine, S. D., Ball, J. W., Plummer, L. N. & Jones, B. F. *U.S.G.S. Water-Resources Investigation Report* 84-4186 (1984).
18. Johnson, N. M., Driscoll, C. T., Eaton, J. S., Likens, G. E. & McDowell, W. H. *Geochim. cosmochim. Acta* **45,** 1421–1437 (1981).
19. Bloom, P. R., McBride, M. B. & Chadbourne, B. *Soil Sci. Soc. Am. J.* **41,** 1068–1073 (1977).
20. Langmuir, D. in *Adsorption from Aqueous Solutions* (ed. Tewai, P. H.) 1–18 (Plenum, New York, 1981).
21. Hargrove, W. L. & Thomas, G. W. *Soil Sci. Soc. Am. J.* **48,** 1458–1460 (1984).
22. Driscoll, C. T. *Int. J. envir. analyt. Chem.* **16,** 267–283 (1984).

A new Miocene hominoid from Kenya

R. E. Leakey & M. G. Leakey

National Museums of Kenya, PO Box 40658, Nairobi, Kenya

A newly discovered early Miocene site in northern Kenya has yielded a partial cranium, several mandibles, isolated teeth and some associated postcranial elements of a large ape that is distinct from other known African and Asian hominoids. The new genus implies a greater complexity in the relationships between the Miocene hominoids and may offer useful evidence concerning the origin of the Asiatic and African large apes. The genus *Sivapithecus* has been reported from East Africa[1] but the new material supports the hypothesis that this genus only occurs beyond Africa[2], although having an African ancestry. The new genus presents a mosaic of characters which suggest an early radiation of Miocene hominoids.

The specimens were found in exposed sediments along the western side of Lake Turkana in northern Kenya in 1985–86 at a small but rich early Miocene locality named Kalodirr (3°20′ N, 35°40′ E). Approximately 10 km south there are other early Miocene fossil-bearing sediments known by several names, including Losodok[3], Moruaret Hill[4], and Moruarot[3,5,6], which probably belong to the same sequence of Miocene deposits. It has not yet been possible to date the locality but we estimate the age may be between 16 and 18 million years based on faunal comparison. Almost all the elements of the mammalian macrofauna recovered are represented at Rusinga (Hiwegi Formation)[7] and the fauna is similar to that from Buluk[1] and Moruarot[6] (Table 1), but three hominoid taxa were recovered all representing new genera. The largest is described here.

Order Primates
Superfamily Hominoidea
Family incertae sedis
Afropithecus gen. nov.

Generic diagnosis. A large hominoid with long narrow palate, distinguished from all known large hominoids by the long distinct muzzle and steeply inclined frontal such that in lateral profile, prosthion, rhinion, glabella and bregma are on almost the same line, and by the corresponding obliquity of the mandibular ramus. It differs from *Sivapithecus, Ramapithecus, Kenya-pithecus, Proconsul* and *Rangwapithecus* in the lingual and buccal basal flare of the upper molars and premolars, particularly P^4, so that the cusp tips tend to be relatively close compared to the cervical width. It differs from *Sivapithecus, Ramapithecus* and *Kenyapithecus* in the buccolingually expanded premolars, the pronounced cingulum bordering the protocone of the upper molars mesially and lingually and the buccolingually broad lower incisors, particularly I_2. It differs from *Rangwapithecus* in the buccolingually expanded upper premolars, the upper P^4 which is shorter mesiodistally than P^3, the less developed lingual cingulum of the upper molars, the procumbent upper incisors and the long superior aspect of the premaxilla. It differs from the similar sized *Proconsul major* in mandibular morphology particularly the lack of development of the superior transverse torus and the steeply sloping subincisive planum.

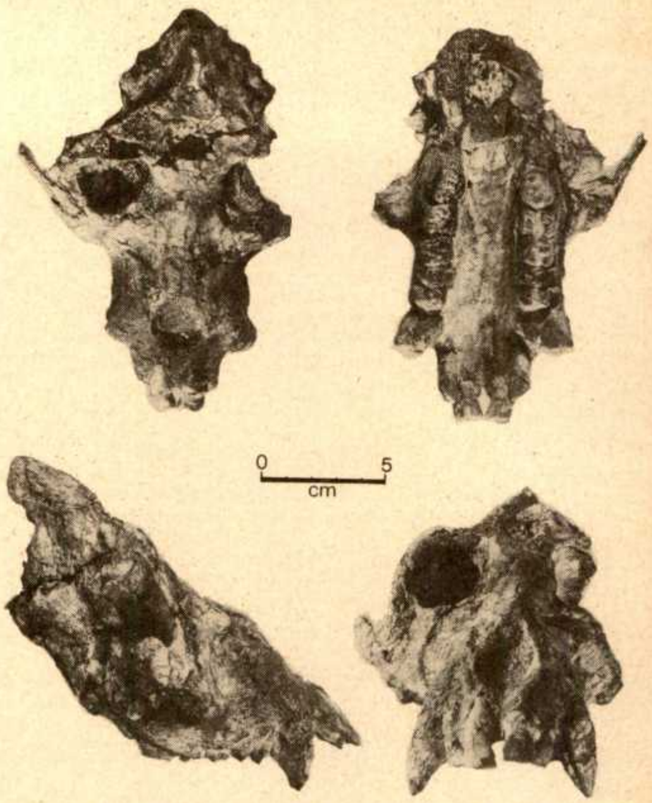


Fig. 1 *Afropithecus turkanensis* cranium, KNM-WK 16999 (type). Above, superior and inferior aspects; below, right lateral and frontal aspects.

Etymology. *Afro* after the continent on which it is found, *pithecos* Greek for ape.

Afropithecus turkanensis sp. nov.

Species diagnosis. As for genus

Type specimen. KNM-WK 16999, facial and frontal region of a cranium with dentition (Fig. 1).

Hypodigm.

KNM-WK 16840	lt. mandible
KNM-WK 16901	lower lt. C, root rt. C, rt. fibula, prox. rt. m/t III, rt. m/t IV, dist. m/tI
KNM-WK 16959	weathered rt. and lt. mandible frags.
KNM-WK 16962	lt. maxilla and premaxilla, rt. C, rt. max. frag., rt. frontal frag.
KNM-WK 17010	associated lower teeth
KNM-WK 17011	rt. mandible frag.
KNM-WK 17012	associated upper teeth
KNM-WK 17013	rt. upper I^1
KNM-WK 17014	mandible frags.
KNM-WK 17016	associated lower teeth, partial shaft rt. ulna
KNM-WK 17021	rt. mandible and associated teeth
KNM-WK 17023	frag. rt. lower C crown
KNM-WK 17024	rt. lower M_2
KNM-WK 17025	lt. lower M_1 frag.
KNM-WK 17026	frag. lower I crown
KNM-WK 17037	lt. lower P_3 crown
KNM-WK 17162	lower I_1 and rt. P_3 frag.
KNM-WK 17168	frag. lower I crown

Etymology. After Lake Turkana.

Table 1 Faunal list

- Mollusca
  - *Melanoides tuberculata*
  - *Limicolaria* sp.
  - *Lanistes carinatus*
  - *Pila ovata*
  - *Etheria elliptica*
- Insecta
  - cf. Agromyzidae
- Pisces
  - Myliobatiformes sp. indet
  - *Polypterus* sp.
  - cf. Bagridae
  - Siluriformes indet.
- Serpentes indet.
- Chelonia
  - Trionychidae indet.
  - Testudinidae indet.
- Crocodilia nov.
- Mammalia
  - *Afropithecus turkanensis**
  - Hominoidea gen. nov. A
  - Hominoidea gen. nov. B
  - *Pronasilio* sp.
  - cf. *Gymnurechinus* sp.
  - *Paraphyomys pigotti*†‡
- Mammalia—continued
  - *Paraphyomys stromeri*†
  - *Paraphyomys* sp.
  - *Elmerimys woodi*†
  - Cricetodontidae indet.
  - *Kenyalogomys minor*†
  - *Anasinopa leakeyi*†‡
  - *Hecubides macrodon*†
  - *Kichechia zamanae*†‡
  - *Hyaenodon* sp.
  - *Platybelodon* sp.*
  - *Prodeinotherium hobleyi**†
  - *Megalohyrax championi**†‡
  - cf. *Aceratherium acutirostratum**†‡
  - *Masritherium aequitoralis*†
  - cf. Hyoboops africanus*†
  - *Kenyasus rusingensis*†‡
  - *Libycochoerus jeanelli*†‡
  - *Lopholistriodon moruoroti**‡
  - *Diamanthohyus africanus**†‡
  - *Canthumeryx sirtensis**†
  - Giraffidae nov.
  - *Dorcatherium parvum*†‡
  - *Dorcatherium pigotti**†‡
  - *Dorcatherium chappuisi**†‡

* Buluk (ref. 6).
† Rusinga, Hiwegi Formation (ref. 5).
‡ Moruarot (ref. 4).

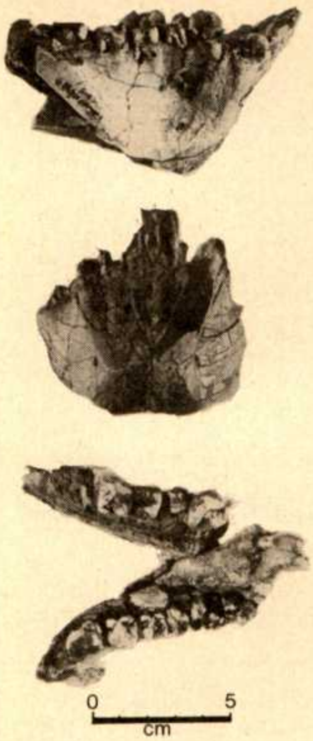


Fig. 2 *Afropithecus turkanensis* mandible, KNM-WK 16840. Above, right lateral aspect; centre, anterior aspect and below, occlusal aspect.

Description. The type specimen, KNM-WK 16999, includes the snout, the facial skeleton, the frontal reaching close to the coronal suture, the pterygoid processes, the sphenoid body, the vomer and the relatively unworn adult dentition.

The palate is shallow, long and narrow and the tooth rows converge posteriorly. The internal nasal aperture is remarkably narrow (15 mm) and high (~18 mm). There is a large (6.5 mm) diastema between the strongly procumbent I^2 and the canine. The enamel surface of the incisors and canines is markedly wrinkled by numerous small vertically running grooves and the occlusal surface of the cheek teeth is covered by numerous wrinkles. The slightly worn central incisors are large and broad while the lateral incisors are relatively small and mediolaterally compressed. The canines are broad and in cross section at the cervix almost circular. Their relative size suggests that the individual was male. The premolars are large, P^3 being larger than P^4. In both the buccal cusp is higher than the lingual and there is marked buccal basal flare. The P^3 has slight lingual basal flare whereas in P^4 this is more marked and the cusp tips are relatively close compared to the basal buccolingual width. $M^1 < M^2 < M^3$ and the molars are broader than they are long (Table 2). On all three molars there is a distinct but short mesial and lingual cingulum extending around the protocone. The enamel cap appears relatively thick judged from the degree of wear before dentine exposure.

The surface bone of the premaxilla is largely replaced by calcite. Its superior aspect is long, forming a gently sloping shelf with a central longitudinal groove. The nasals are long and narrow (13 mm across the widest point) but the tips are lost close to the margin of the small oval pyriform aperture. The left side of the maxilla is best preserved. There is a distinct jugum and posterior to it a small postcanine fossa. On either side of the nasals, anterior to the marked lacrimal foramina, the maxilla is distinctly hollowed. The zygomatic process is deep below the orbit (37 mm) and the inferior border swings markedly upwards from the masseteric tubercle towards the temporal, a striking feature probably related to an unusually high position of the cranial vault. The right orbit is virtually complete and relatively undistorted. It is higher than it is wide and has an asymmetrical outline; the greatest depth (30 mm) is lateral to its longitudinal axis and the greatest width (25 mm) is close to the superior margin. The supraorbital tori are very thin and there is a small but distinct supraorbital notch. The interorbital distance is large (24.5 mm). The marked temporal lines swing sharply towards the midline forming a frontal trigon which rises steeply towards the apex, approximately 52 mm posterior to glabella. There is a distinct, mostly midline frontal sinus, seen through the breaks in the region immediately above glabella. Glabella does not project anterior to the orbits and is weakly developed. In lateral profile, prosthion, rhinion, glabella and bregma are almost in the same line, a feature not observed in any other fossil or living large hominoid.

KNM-WK 16840 (Fig. 2) is the most complete of the mandibular specimens. The body is deep and robust and the tooth row relatively short. The symphysis is deep (60 mm) and the internal surface slopes steeply down to the genial fossa with little development of the transverse tori. The superior anterior aspect is the widest because of the lateral flare of the canines. The narrowest point is at a level with the mental foramen and below this, it widens towards the prominent mental tubercles. The mandibular body is deepest anteriorly (54 mm at the posterior canine alveolar margin) and becomes slightly shallower towards the posterior (44 mm below M2). There is a distinct mandibular fossa. There is a clearly defined and very large digastric fossa. One specimen, KNM-WK 17021, preserves details of the ramus which is set at an oblique angle to the body. There is a slight inferior projection of the mandibular angle.

Several specimens provide details of the lower dentition at various stages of wear (Table 2). The lower incisors are all marked by varying degrees of mesial curvature from the root tip to the crown and the roots are broad labiolingually, particularly I_2. The P_3s are relatively large, rather narrow and sectorial with a slight extension of the enamel onto the buccal face of the mesial root. The P_4s are small compared to the P_3s, low crowned and bicuspid. The molars are slightly inflated buccally and on one specimen, KNM-WK 17024 there is evidence of a slight buccal cingulum. M_3 has a small tuberculum sextum. $M_1 < M_2 < M_3$.

Afropithecus displays characters typical of a variety of Miocene hominoids combined in a single taxon that should also include the large hominoid recently described from Buluk[1]. We

Table 2 *Afropithecus* tooth dimensions (mm)

	Upper			Lower		
Tooth	Acc. no.	mesiodistal	buccolingual	Acc. no.	mediodistal	buccolingual
I1	16999	11.6	9.2	16840	(5.2)	8.0
				17016*	(4.5)	8.4
				17021*	5.1	8.6
				17162	8.1	5.5
I2	16999	7.6	9.7	16840	(8.0)	10.0
				17010	6.5	9.5
				17016*	6.0	10.5
				17021	6.0	10.8
C	16999	18.0	17.4	16901	12.2	15.4
				17014*	—	(15.0)
				17016	12.0	16.8
P3	16999	9.0	14.6	17016	8.9	(14.3)
P4	16999	7.8	14.9	16840	(15.0)	—
	17012	6.5	12.0	17010	8.0	7.8
				17014*	(12.5)	8.3
				17016*	8.0	(10.0)
				17021	(8.5)	9.6
M1	16999	(11.5)	(13.0)	17016	(10.3)	—
	17012	9.0	11.2	17021	(10.5)	—
M2	16999	(12.0)	(14.0)	17010	12.5	(10.5)
	17012	10.0	11.5	17014	12.0	11.0
				17016	12.4	11.5
				17021	(13.0)	—
				17024*	12.0	10.5
M3	16999	(12.5)	(13.0)	17010*	14.3	13.0
	17012	10.4	10.5	17014*	14.0	—
				17016	14.0	11.0
				17021*	(14.5)	(11.5)
C-M3	16999	69	—	—	—	—
P3-M3	16999	52	—	—	—	—
M1-M3	16999	36	—	—	—	—

Left tooth dimensions are given except where * indicates right tooth dimensions.

now believe that *Sivapithecus*[8,9] is restricted to Asia including the Siwaliks, Indo-Pakistan[2]. A careful comparison of the mandibular morphology of *Proconsul major* as represented by the material known from Songhor and Koru[10] differs from *Afropithecus* in the strong development of a superior transverse torus and lack of an inferior torus, the long shallow alveolar subincisive planum and the marked increase in size of the lower molars from M_1 to M_3, not seen in *Afropithecus.* The maxillary morphology of *P. major* is poorly represented. The only reasonably complete specimen that may represent this taxon is the Moroto palate and associated facial fragments. Its assignment [10-13] to *P. major* is disputed[14,15] but when the Moroto palate is compared to KNM-WK 16999, the superior aspect of the premaxilla is shorter, the breadth of the nasal aperture is greater, P_4 is less buccolingually expanded and there is no well developed lingual cingulum on the molars.

One weathered maxillary fragment from Moruorot, KNM-LS 7, previously referred to *P. major*[10] was recently designated the type of *Proconsul* (*Xenopithecus*) *hamiltoni*[16]. The two eroded and worn molar crowns[17,18] are both elongated buccolingually whereas those of *Afropithecus* are more square. Together with the limited details of cusp morphology, this suggests that this specimen is not *Afropithecus.* In fact the specimen is so poor that we suggest that the name *P.* (*Xenopithecus*) *hamiltoni* be abandoned.

Afropithecus has been compared with the poorly represented African genus *Kenyapithecus.* Little is known of the maxillary morphology of either the later species, *K. wickeri,* or the earlier species, *K. africanus,* but both differ from *Afropithecus* in the dentition, particularly the lack of buccal basal flare on the molars and premolars, the premolars that are not buccolingually expanded and the molars that lack a marked cingulum bordering the protocone. Similarities are apparent in the root of the zygomatic which leaves the maxilla above M^1 and the postcanine fossa of *K. wickeri.* The mandible of *K. wickeri* is known but differs from *Afropithecus* in having a relatively shallow body with a shallow rather gently sloping subincisive planum.

The phyletic status of the large hominoids from the Miocene of Europe is not agreed[19-21] but we have considered the published descriptions and made direct comparisons between *Afropithecus* and casts available to us of the European fossils and we are satisfied that *Afropithecus* is quite distinct from the European species of *Dryopithecus*[22], *Ouranopithecus*[21,23], *Rudapithecus* and *Bodvapithecus*[22].

Afropithecus resembles the specimen GSP 15000 from the Siwaliks[24] attributed to *Sivapithecus sivalensis* (formally *S. indicus*[20]) in the postcanine fossa, the large procumbent central incisors, the relatively small asymmetrical lateral incisors, the long superior aspect of the premaxilla, the small suprorbital tori and the high hafting of the cranial vault onto the face. It differs in the distinct snout and the extraordinary profile with the almost linear alignment of prosthion, rhinion and glabella. These differences also apply when a comparison is made with the known large hominoids from Lufeng[25]. The Lufeng specimens resemble *Afropithecus,* however, in the conformation of the temporal lines and the wide interorbital distance. Details of sub-nasal morphology which have been used as a basis for discriminating between various hominoids[26,27] cannot be discerned on KNM-WK 16999.

In conclusion, the mosaic of characters represented in the new genus, which precludes its inclusion in any of the Miocene hominoid families previously defined, suggests a greater complexity in the early Miocene than was previously apparent. The absence of *Afropithecus* from the richly fossiliferous localities of western Kenya where fossil primates are common may indicate ecological or temporal factors. Further work at Kalodirr and other Miocene sites in northern Kenya will investigate these questions.

We thank the Government of Kenya and the Governors of the National Museums of Kenya. The research was financed by the National Museums of Kenya and contributions from The Maytag Family Foundation, Mr and Mrs L. Newitt and the National Geographic Society of Washington, D.C. We also thank Bw Kamoya Kimeu and his prospecting team, particularly Bw Wambua Mangao who found the type specimen, Bw Joseph Mataba who prepared the majority of the specimens, Dr L.

Jacobs, P. Kat and Ms K. Stewart who identified the microfauna, molluscs and fish respectively, and Dr Alan Walker who assisted in many ways.

Received 25 June; accepted 15 September 1986.

1. Leakey, R. E. F. & Walker, A. *Nature* **318,** 173-175 (1985).
2. Kelley, J. & Pilbeam, D. *Am. J. phys. Anthrop.* **66,** 188 (1985).
3. Arambourg, C. *Ann. Paleont.* **22,** 121-146 (1933).
4. Deraniyagala, P. E. P. *Spolia zeylan.* **25,** 1-42 (1948).
5. Walsh, J. & Dodson, R. G. *Geol. Surv. Kenya* **82,** 1-42 (1969).
6. Madden, C. T. *Paleobios* **14,** 1-12 (1972).
7. Pickford, M. *Kenya Palaeont. Gaz.* **1** (National Museums of Kenya, Nairobi, 1984).
8. Pilbeam, D., Smith, R. *Mem. Geol. Surv. Pakistan* **11,** 1-13 (1981).
9. Pilbeam, D. R., Rose, M. D., Badgley, C. & Lipschutz, B. *Postilla* **181,** 1-94 (Peabody Mus. Nat. Hist., Yale University, 1980).
10. Andrews, P. J. *Bull. Br. Mus. (nat. Hist). Geol.* **30,** 85-244 (1978).
11. Simons, E. L., Andrews, P. & Pilbeam, D. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 120-146 (Harvard University Press, Cambridge, 1978).
12. Allbrook, D. & Bishop, W. W. *Nature* **197,** 1187-1190 (1963).
13. Pilbeam, D. *Bull. Peabody Mus. nat. Hist.* **31,** 1-185 (1969).
14. Leakey, L. S. B. in *Classification and Human Evolution* (ed. Washburn, S. L.) 32-49 (Viking Fund Publications in Anthropology, No. 37, 1963).
15. Martin, L. *J. hum. Evol.* **10,** 139-150 (1981).
16. Madden, C. T. *Primates* **21,** 241-252 (1980).
17. Hopwood, A. T. *Ann. Mag. nat. Hist.* (Ser. 10) **11,** 96-98 (1933).
18. Le Gros Clark, W. E. & Leakey, L. S. B. *Fossil Mammals of Africa* (Br. Mus. Nat. Hist.) **1,** 1-117 (1951).
19. Kay, R. F. & Simons, E. L. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 577-624 (Plenum, New York, 1983).
20. Bonis, L. de & Melentis, J. *C.r. hebd. Séanc. Acad. Sci., Paris* (*D*) **284,** 1393-1396 (1977).
21. Simons, E. L. & Pilbeam, D. R. *Folia primat.* **3,** 81-152 (1965).
22. Kretzoi, M. *Nature* **257,** 578-581 (1975).
23. Martin, L. & Andrews, P. *Cour. Forsch. Inst. Senckenberg.* 6925-6940 (1984).
24. Pilbeam, D. *Nature* **295,** 232-234 (1982).
25. Wu Rukang *Cour. Forsch. Inst. Senckenberg* **69,** 41-48 (1984).
26. Ward, S. C., Kimbel, W. H. & Pilbeam, D. *Am. J. phys. Anthrop.* **61,** 157-171 (1983).
27. Ward, S. C. & Pilbeam, D. R. in *New Interpretations of Ape and Human Ancestry* (eds Ciochon, R. L. & Corruccini, R. S.) 211-238 (Plenum, New York, 1983).

A second new Miocene hominoid from Kenya

R. E. Leakey & M. G. Leakey

National Museums of Kenya, PO Box 40658, Nairobi, Kenya

A recently discovered early Miocene site in northern Kenya (see preceding paper[1]) has yielded a number of important fossil hominoid specimens including a partial cranium, mandible and postcranial remains of a medium-sized ape that is distinct from all previously reported genera. The relationship of this taxon to other known hominoids is not clear. A new genus of large hominoid was found at the same site[1]. The recovery of two new genera of Miocene hominoids which are apparently not represented at the well known and rich Miocene localities of western Kenya may indicate that a different ecology is being sampled.

The specimens were found in 1985-6 at a new Miocene locality named Kalodirr on the western shores of Lake Turkana. The site has not been dated but a faunal comparison suggests it may be between 16 and 18 million years[1].

Order Primates
Superfamily Hominoidea
Family incertae sedis
Turkanapithecus gen. nov.

Generic diagnosis. A shortfaced small hominoid approximately the size of a *Colobus polykomos*. It is distinguished from all known hominoids by the upper M^2 s, which have an additional cuspule between the pronounced and beaded mesial and lingual cingula, and the upper molars and P^4 which have small additional cuspules associated with a small buccal cingulum bordering the paracone. It is distinguished from the Oreopithecidae by the relatively low relief of the dental occlusal surface, upper molars not mesiodistally elongated and lower molars not long and narrow, without buccolingual waisting and lacking a centroconid. It is distinguished from the larger *Rangwapithecus gordoni* by the upper M^3 approximately the same size or smaller than M^2 and the upper P^4 smaller than P^3. It is distinguished from the slightly larger *Proconsul africanus* by the distinct snout, the wide interorbital distance, the relatively thick supraorbital tori, the broad nasals, the relatively broad and low mandibular ramus and the upper premolars and molars not buccolingually expanded. It is distinguished from the similar sized *Pliopithecus vindobonensis* by the distinct snout, the long nasals, and the zygomatic process of the maxilla approximately vertical with no protruding inferior orbital margin.

Table 1 *Turkanapithecus* tooth dimensions (mm)

	Acc. no.	Tooth	mesiodistal	buccolingual
Upper	KNM-WK 16950	C*	10.0	7.9
		P^3	6.8	8.0
	KNM-WK 16957	P^3	5.9	7.7
	KNM-WK 16950	P^4	5.5	—
		M^1	7.0	7.5
		M^{2*}	8.6	9.2
		M^3	8.5	8.9
		$C-M^{3*}$	44.0	—
		P^3-M^3	35.2	—
		M^1-M^3	24.0	—
Lower	KNM-WK 16950	M_2	9.2	7.0
		M_3*	9.2	7.5

Left tooth dimensions are given except where * denotes right tooth dimensions.

Etymology. After Lake Turkana and *pithecos*, Greek for ape.

Turkanapithecus kalakolensis sp. nov.

Species diagnosis. As for genus.

Type specimen. KNM-WK 16950A and B, a partial cranium and mandible (Figs 1 and 2).

Hypodigm.

KNM-WK 16950I-X postcranial elements probably associated with the type specimen: right femur, proximal and distal left radius, left metacarpal I and fragments of phalanges, rib and vertebrae

KNM-WK 16957 fragment of maxilla.

Etymology. After the River Kalakol of which the Kalodirr is a tributary.

Description: The type cranium, KNM-WK 16950, which is slightly smaller than the female *Proconsul africanus* skull (KNM-RU 7290) is reasonably complete and includes the facial skeleton, partial calvaria and examples of all the permanent upper teeth except the incisors. The associated mandible preserves much of the left ramus, the mandibular body and the left M_2 and right M_3. A complete right femur, fragments of proximal and distal left radius, ribs, vertebrae and foot bones are most probably associated with the cranium and mandible but there is a slight element of uncertainty concerning this association, so they are not designated as part of the type material.

The palate is rather shallow and narrow with the tooth rows converging towards the back. A crack, 2.2 mm wide, runs along the centre of the palate making it appear broader than it was. Judged from the position of the back of the palate and the length of the remaining right pterygoid, the internal nasal aperture was high and narrow. The teeth (Table 1) are moderately to heavily worn. The right upper canine is relatively large suggesting the individual was male. The right P^4 is missing and was lost in life causing some malocclusion; the wear on the right P^3 is greater than on either the left P^3 or the left P^4. P^3 is significantly larger than P^4, both are wide buccolingually. The upper molars are rather square in outline but wider than they are long. $M^1 < M^3 < M^2$. The morphology of the M^1s is obscured by wear, dentine being exposed over almost the entire occlusal surface. The M^2s are distinctive, having marked, beaded mesial and lingual cingula with an unusual accessory cuspule between them. The M^3s

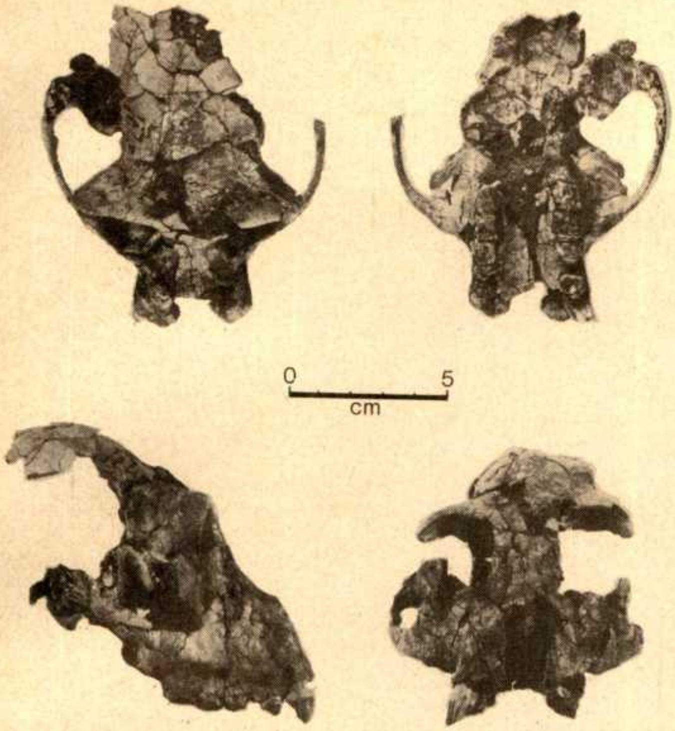


Fig. 1 *Turkanapithecus kalakolensis* cranium, KNM-WK 16950A (type). Above, superior and inferior aspects; below, right lateral and frontal aspects.

are more worn than the M^2s, probably due to the apparent high set of the lower M_3s. P^4 to M^3 have small buccal cingula bordering the paracone each with an additional cuspule variably developed.

The premaxilla is largely missing. The pyriform aperture was of moderate size and oval (approx. 17 mm wide). The nasal bones are large, being long (maximum length 27 mm) but also fairly broad (maximum width 14 mm across both). The right maxilla is the best preserved; there is a slight canine jugum but no postcanine fossa. The lower border of the root of the zygomatic process is about 7 mm above M^1. The zygomatic process is relatively deep below the orbit with the lower border being close to vertical beneath the lower orbital margin. The orbits are incomplete and slightly distorted; they were probably almost circular but slightly higher (approx 24 mm) than wide (approx 22 mm). There is only the slightest supraorbital notch and no development of a supraorbital sulcus. The interorbital distance is large (21 mm). In the region of glabella the bone is crushed and fractured exposing a fairly well developed frontal sinus. The frontal is distorted on the left side lateral to the rather indistinct temporal line. The right temporal line is strongly developed passing initially towards the midline almost parallel with the orbital margin before curving sharply towards the posterior. Little of the parietals is preserved. The occipital area is damaged but there is a fragment that includes the posterolateral margins of the foramen magnum with most of both condyles and a small portion of bone extending to the nuchal crest. The temporal bone is partially preserved on the right and includes the mandibular fossa, the postglenoid process and the zygomatic process which attaches to the zygomatic.

The mandible is relatively complete particularly on the left side which lacks only some bone in the regions of the angle of the mandible, the coronoid process and the canine alveolus. The mandible is gracile and deeper at the symphysis (28 mm) than below M_2 (21 mm). The internal surface of the symphysis slopes evenly down as a narrow, slightly concave subincisive planum. There is virtually no development of a superior or inferior transverse torus. A small digastric fossa is clearly defined on the right. The anterior root of the broad ascending ramus originates close to M_2. The mandibular condyle is relatively large and rounded being broad mediolaterally and wide anteroposteriorly. Only two lower teeth were recovered, the moderately worn left M_2 and the right M_3. The lingual cusps are positioned slightly distal to the buccal and the talonid basin is large. On M_2, both the mesial and distal foveae are clearly defined and the hypoconulid was probably large.

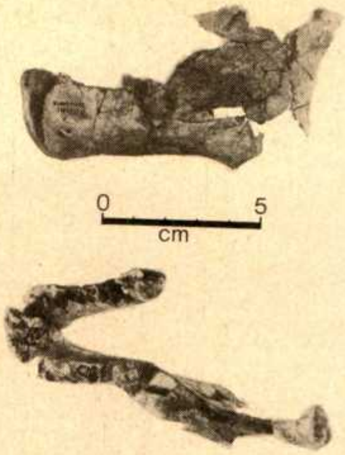


Fig. 2 *Turkanapithecus kalakolensis* mandible, KNM-WK 16950B (type). Above, left lateral and below, occlusal aspects.

The phylogenetic relationship of *Turkanapithecus* to other known fossil hominoids is not clear. The cranial morphology has been compared with the only other similar sized African Miocene hominoid for which a skull is known, *Proconsul africanus*[2,3] and there are important differences. *Turkanapithecus* is distinguished by the distinct and rather 'square' maxilla, the larger pyriform aperture, the longer nasals which widen at either end, the greater depth of the zygomatic process of the maxilla beneath the orbit, the root of the zygomatic process above M^1 not M^2, the wider interorbital distance, the distinct supraorbital tori and the rather flat supraglabella region. The mandibular ramus is lower and wider anteroposteriorly and the condyle is knoblike and not elongated mediolaterally as in *P. africanus*. Comparison with the similar sized cranium of *Pliopithecus vindobonensis* from the Middle Miocene of Europe[4,5] also shows differences in the 'square' maxilla of *Turkanapithecus*, the long nasals and the zygomatic process of the maxilla roughly vertical with no protruding inferior orbital margin. The dentition differs from both *Pliopithecus* and *Proconsul* in the features mentioned in the diagnosis.

The dentition has been compared with some isolated teeth from Maboko which have recently been attributed to the previously monospecific family, the Oreopithecidae[6] and included in a new genus to which *Rangwapithecus vancouveringi* is also referred[7]. They are considered to have some affinities with *Rangwapithecus gordoni*. These teeth may be readily distinguished from *Turkanapithecus* in the high cusped elongated molars. Szalay and Delson[8] suggest that *Oreopithecus* has cercopithecoid affinities on the basis of a number of shared derived characters of the dentition. These features are those which distinguish *Oreopithecus* from *Turkanapithecus*.

Unfortunately little is known of the cranial morphology of the Miocene small bodied apes. *Limnopithecus*, *Dendropithecus* and *Micropithecus*[9,10] but the dentition clearly separates these genera from *Turkanapithecus*. Although ecological differences may be a factor, it is puzzling, if our age estimates are correct, that *Turkanapitheus* has not been recorded previously from any of the known Miocene sites in western Kenya where fossil apes are common. This relatively complete material from Kalodirr is important in the new evidence it provides of greater diversity in the Miocene Hominoidea.

We thank the Government of Kenya and the Governors of the National Museums of Kenya. The research was financed by the National Museums of Kenya and contributions from The Maytag Family Foundation, Mr and Mrs L. Newitt and the National Geographic Society of Washington, D.C. We thank Bw. Kamoya Kimeu and his prospecting team who discovered all the fossils described and particularly Bw. Wambua Mangao who found the type specimen. Bw. Joseph Mataba prepared the specimens and Dr Alan Walker assisted in many ways.

Received 25 June; accepted 15 September 1986.

1. Leakey, R. E. F. & Leakey, M. G. *Nature* **324,** 143-146 (1986).
2. Le Gros Clark, W. E. & Leakey, L. S. B. *Fossil Mammals of Africa* (Br. Mus. Nat. Hist) **1,** 1-117 (1951).
3. Walker, A., Falk, D., Smith, S. & Pickford, M. *Nature* **305,** 525-527 (1983).
4. Hurzeler, J. *A. Paleont.* **40,** 5-63 (1954).
5. Zapfe, H. *Mem. Suisses Paleontl.* **78,** 5-293 (1961).
6. Harrison, T. *Am. J. phys. Anthrop.* **66,** 180 (1985).
7. Harrison, T. *Am. J. phys. Anthrop.* (in the press).
8. Szalay, F. S. and Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
9. Harrison, T. thesis, London Univ. (1982).
10. Fleagle, J. and Simons, E. *Am. J. phys. Anthrop.* **49,** 427-440 (1978).

Vegetation, climatic and floral changes at the Cretaceous-Tertiary boundary

Jack A. Wolfe & Garland R. Upchurch, Jr

US Geological Survey, MS-919, Box 25046, Federal Center,
Denver, Colorado 80225, USA

The western interior of North America has the only known non-marine sections that contain the iridium-rich clay interpreted as the Cretaceous–Tertiary (K–T) boundary[1–7]. Because vegetation and climate can be directly inferred from physiognomy of leaves[8–15] and because leaf species typically represent low taxonomic categories, studies of leaf floras in these sections provide data on the effects of a terminal Cretaceous event on the land flora, vegetation and climate. A previous study based on detailed sampling of leaves and their dispersed cuticle[16] in the Raton Basin provides a framework for interpretation of other leaf sequences over 20 degrees of latitude. We conclude that at the boundary there were: (1) high levels of extinction in the south and low levels in the north; (2) major ecological disruption followed by long-term vegetational changes that mimicked normal ecological succession; (3) a major increase in precipitation; and (4) a brief, low-temperature excursion, which supports models of an 'impact winter'.

Latest Cretaceous (late Maestrichtian, equivalent to Lancian; Phase 1 (ref. 16)) vegetation from palaeolatitude[17] 46° to >56 °N (Fig. 1) was dominantly broad-leaved evergreen (Table 1). In the south small leaf size, paucity of drip tips and thick cuticles with cutinized hairs on both surfaces indicate dry, open-canopy forest[16,18]. In the Raton sequence (Fig. 2) dispersed-cuticle assemblages from the base of the formation to <1 cm below the K–T boundary show no physiognomic changes that indicate significant change in climate. Although leaf size increased into higher latitudes, drip tips were uncommon, suggesting subhumid forest. At 66° N, vegetation was large-leaved deciduous forest that resulted from the polar light regime[19].

Broad-leaved decidous plants were not diverse in mid-latitude Lancian vegetation (Fig. 2), as is true for modern Southern Hemisphere vegetation; thus we use the Southern Hemisphere ratio between the proportion of entire-margined species (EM) and mean annual temperature[10]. The 20 °C isotherm (megathermal/mesothermal boundary) falls at 50°–53° N, consistent with major floristic differences in megafloras[15] and palynofloras[20]. The latitudinal temperature gradient was ~0.3 °C per degree of latitude (Fig. 3), a gradient also obtained for the North American Campanian[18]. Comparisons of early to late Lancian assemblages in the same areas indicate no significant change in EM and thus no change in temperature.

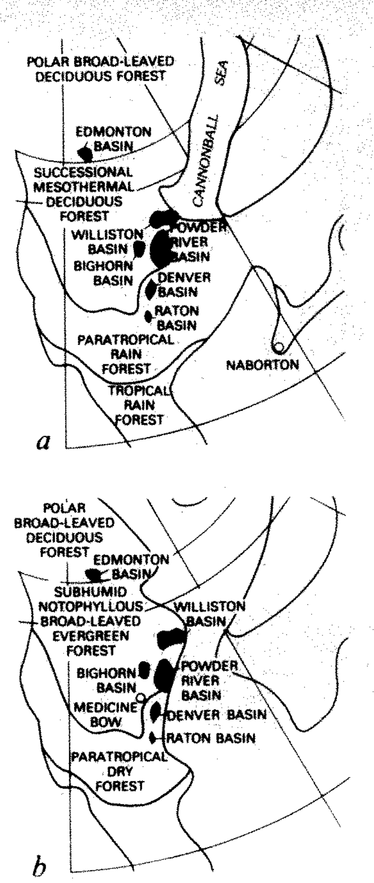


Fig. 1 Maps of the western interior of North America during *a*, the early Palaeocene and *b*, the latest Cretaceous showing major depositional basins, floral assemblages and inferred vegetation. Mesothermal vegetation includes the southern part of polar broad-leaved deciduous forest, subhumid notophyllous broad-leaved evergreen forest and successional mesothermal deciduous forest. Megathermal vegetation includes paratropical dry forest, paratropical rain forest and tropical rain forest (see refs 10, 19). Throughout the Late Cretaceous and early Tertiary the southern margin of polar broad-leaved deciduous forest fell between paleolatitudes 65° and 70° N (refs 18, 19). Continental position adapted from ref. 17); mid-Maestrichtian and early Palaeocene epicontinental seas adapted from ref. 26.

Palaeocene assemblages <3 m above the K–T boundary (Phases 2 and 3) indicate ecological disruption and mass kill. Assemblages of Phase 2, which includes the palynological fern spike[5], are dominated by herbaceous monocots and/or ferns, which persist into Phase 3. Phase 3 assemblages are low diversity and have leaves of early successional morphology[12]. Included are narrow-leaved taxa that typically occurred in disturbed, fluviatile habitats during the Lancian, but during early succession following the K–T boundary opportunistically invaded swamp habitats. Temperature inferences are difficult for such a limited flora, but the large leaf size suggests abundant moisture.

All early Palaeocene assemblages of Phases 4 and 5 have large leaf size with no latitudinal increase. Increase in leaf size over Lancian assemblages from 46° to 56° N indicates increased precipitation, as does an increase in the proportion of drip-tipped species and a change in cuticular physiognomy in the south. In the Raton sequence this change occurs at the K–T boundary[16]. This increase would lead to high-biomass vegetation and high water tables and was possibly the major causative factor in the widespread initiation of peat deposition that marks the boundary in the northern High Plains. Persistence of high precipitation indicates that the climate was pushed into a new, metastable mode. Initiation of megathermal rain forest environments had major biotic consequences[18,21].

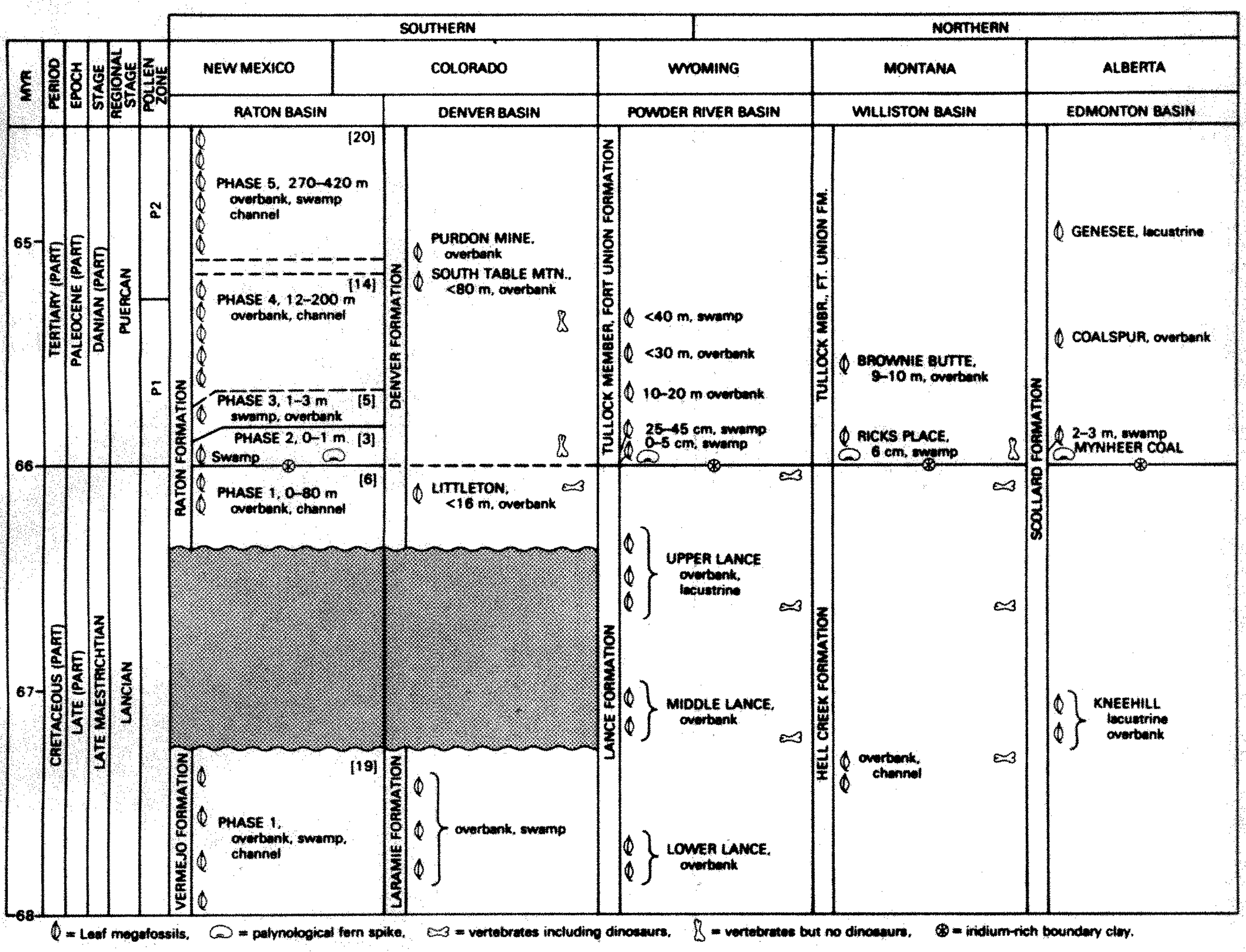


Fig. 2 Stratigraphical and depositional occurrences of some plant megafossil assemblages relative to the Cretaceous-Tertiary boundary in the western interior. Estimates of absolute ages are based on some radiometric and some palaeomagnetic data[34,35]. Palynological data from refs 5, 36–38; stratigraphical data from refs 16, 39 and unpublished data; Physiognomic and floristic data from refs 16, 40–44 and unpublished data. Following some floral symbols are measured or estimated stratigraphical interval from the boundary. Numbers in brackets in the Raton Basin column are numbers of individual localities.

Diversity in Phases 4 and 5 increased over Phases 2 and 3 but was still lower than in Phase 1. Percentages of entire-margined and evergreen species sharply decline northward compared to the Lancian. During the early Palaeocene, megathermal forest extended to 50°–53° N (as in the Lancian) in the western interior but was anomalously bounded on the north by broad-leaved deciduous forest; mesothermal, notophyllous, evergreen forest is unknown. Anomalous deciduousness characterizes most extant mesothermal vegetation of southeastern North America, resulting from prolonged, intense freezes[10]. If such climates characterized the northern High Plains early Palaeocene there would not be a diverse fauna of large ectotherms.

However, occurrence in the northern High Plains of large and presumably ectothermic reptiles (turtles, champsosaurs, and crocodilians[22], some >3 m long; W. S. Bartels, personal communication) indicates that earliest Palaeocene temperatures were, as in the Lancian, mesothermal. The Raton data also indicate the same temperature during the Lancian as during the early Palaeocene, and the megathermal/mesothermal boundary was unchanged. These data make previous temperature estimates[23,24] for the northern high plains early Paleocene, based on deciduousness and low EM, suspect.

Thin leaves, mostly from deciduous plants, typically have toothed margins[12]. Thus if vegetation is anomalously deciduous, a low EM relative to temperature results. This phenomenon is apparent in the western interior early Palaeocene: if the EMs for the leaf assemblages give a valid measure of temperature, the latitudinal temperture gradient from New Mexico to Montana would be 1.2 °C per degree of latitude (Fig. 3). This gradient, which is one-third higher than the present gradient from the southern high plains to Alberta, would produce a mean annual temperature of −2 °C in central Alberta and a polar temperature of −28 °C. Such low temperatures are clearly contradicted by the climatic implications of the Genesee and other northern assemblages[25] and would be unlikely because of the transgressing early Palaeocene Cannonball Sea[26]. Deciduous immigrants in the Denver and elimination of most broad-leaved evergreens in northern Phase 5 assemblages skew the leaf-margin analysis. In Raton Phase 5 vegetation northern deciduous taxa were few and did not affect EM[16].

Oxygen-isotope data[27] suggest a temperature increase from the early into the late Maestrichtian, and analyses of North American Maestrichtian floras[18] parallel marine data. No climatic change can be detected between early and late Lancian assemblages in the late Maestrichtian. Lancian temperature is indistinguishable from early Palaeocene temperature.

Palynological interpretations of the K–T boundary in the western interior suggest little extinction, with most taxa persisting in refugia[5], but are complicated by the facts that (1) palynofloras rarely contain Magnoliidae pollen; and (2) a palynomorph

Table 1 Foliar physiognomy, inferred vegetation and climate and significant taxa of Cretaceous–Tertiary boundary assemblages in the western interior

				Dicotyledons							
Area	Assemblage	Phase	Number of species	Number of species	entire margin (EM) (%)	Leaf-size index* (%)	Coriaceous (evergreen) (%)	Evergreen drip tipped (%)	Inferred vegetation†	Inferred climate‡	Significant taxa
Central Alberta (Edmonton Basin)	Genesee	4–5	26	21	19	73	5	100	PBLDF	W Ms	*Metasequoia*, trochodendroids, platanoid, *Eucommia*, *Acer*, *Averrhoites*, Hamameloideae. Tiliaceae, Iodeae, Nyctaginaceae
	Coalspur	3–4	3	2	(0)	75	0	—	SPBLDF	W Ms?	*Metasequoia*, trochodendroid, Hamameloideae
	Mynheer coal (above)	3	8	3	(33)	50	0	—	S	W Ms?	*Metasequoia*, *Glyptostrobus*, *Thuites*, trochodendroid, flacourt, *Paranymphaea*
	Kneehills, Brazeau	1	21	11	(27)	78	0	—	PBLDF	M Ms	Ginkgophytes, *Glyptostrobus*, *Metasequoia*, *Nilssonia*, trochodendroids, Hamameloideae, *Fagopsis*, '*Acer*', *Thuites*, flacourt
Eastern Montana (Williston Basin)	Brownie Butte	3–4	15	11	(27)	73	0	0	SMDF	W Ms	*Metasequoia*, *Glyptostrobus*, *Thuites*, *Paranymphaea*, trochodendroid, platanoid, Hamameloideae, Tiliaceae
	Ricks Place	2–3	4	1	(100)	50	0	0	S	W Ms	aff. *Stenochlaena*, aff. *Averrhoites*, herbaceous monocot, leafy bryophyte
	Hell Creek	1	27	21	62	53	76	19	NBLEF	S Ms	Cycad, lauraleans, flacourt, '*Dryophyllum*', Hamameloideae, *Ginkgo*, trochodendroid
Eastern Wyoming (Powder River Basin)	Lower Tullock near Newcastle	3–4	22	12	(33)	—	19	100	SMDF	W Ms	Ferns, *Metasequoia*, *Glyptostrobus*, *Thuites*, *Paranymphaea*, aff. *Averrhoites*, trochodendroid, '*Myrtophyllum*', '*Carya*'
	Lower Tullock near Lance Creek & Glenrock	3–4	17	14	(43)	—	36	60	SMDF/SPRF	W Ms/W Mg	Cycad, *Paranymphaea*, laurel, *Corylites*, '*Carya*', '*Myrtophyllum*', platanoid, trochodendroid, menisperm, palm, Hamameloideae
	Lower Tullock (+25 cm)	3	5	4	(75)	75	75	33	S	W Ms/W Mg	aff. '*Cissites*' *panduratus*, aff. *Averrhoites*, celastracean, aff. *Stenochlaena*
	Lowermost Tullock	2	1	0	—	—	—	—	S	?	Herbaceous monocot
	Upper Lance	1	33	24	54	36	76	4	NBLEF	S Ms	Araucarian, lauraleans, '*Dryophyllum*',
	Middle Lance	1	20	17	(56)	39	82	12	NBLEF	S Ms	flacourt, euphorb, '*Myrtophyllum*', trochodroid,
	Lower Lance	1	36	28	57	35	79	0	NBLEF	S Ms	palm
Southern Wyoming	Medicine Bow	1	56	42	67	55	71	13	NBLEF/PDF	S Ms/S Mg	*Geinitzia*, lauraleans, '*Dryophyllum*', Rhamnaceae, '*Myrtophyllum*', trochodendroid
Northeastern Colorado (Denver Basin)	South Table Mount. & Purdon Mine	4–5	56	45	62	74	74	44	SPRF	W Mg	Lauraleans, Rhamnaceae, euphorb, '*Myrtophyllum*', aff. *Grewia*, Hamameloideae, palm
	Littleton	1	54	50	71	57	86	10	PDF	S Mg	Diverse lauraleans, diverse euphorbs Rhamnaceae, palm [under study]
	Laramie	1	60	55	71	52	82	11	PDF	S Mg	Diverse lauraleans, euphorbs, *Geinitizia*, '*Myrtophyllum*', palm [needs restudy]
Southeastern Colorado and Northeastern New Mexico (Raton Basin)	Phase 5	5	50	45	74	68	85	55	PRF/SPRF	W Mg	Lauraleans, '*Magnolia*', euphorb, aff. Heliocarpus, aff. *Grewia*, '*Carya*', palms
	Phase 4	4	26	25	71	72	90	42	SPRF	W Mg	'*Magnolia*', '*Cissites*', celastracean, lauralean, platanoid, *Averrhoites*, palm
	Phase 3	3	8	5	(80)	70	100	40	S	W Mg	'*Cissites*', celastracean, lauralean, ferns, palm
	Phase 2	2	0	—	—	—	—	—	S	W Mg	aff. *Stenochlaena*, leafy byrophyte
	Phase 1 (Raton)	1	47	43	72	34	89	9	PDF	S Mg	Diverse lauraleans, dierse euphorbs, palms, *Geinitzia*
	Phase 1 (Vermejo)	1	82	63	71	34	85	9	PDF	S Mg	Diverse lauraleans, diverse euphorbs, palms, *Geinitzia*, araucarian, '*Mytrophyllum*'
Louisiana	Naborton	>5	47	44	83	75	91	63	TRF	W Mg	'*Magnolia*', lauraleans [needs restudy]

Numbers in parentheses are based on small samples and may not be reliable. Significant taxa are primarily those numerically common in the samples. Scientific names which are not correct, but for which no alternative has been proposed, are shown in quotation marks. Aff. indicates related taxon.

* A few previous collections were not examined; from taxa listed most physiognomic parameters can be calculated, but not leaf size index.

† NBLEF, notophyllous broad-leaved evergreen forest; PBLDF, polar broad-leaved deciduous forest; PDF, paratropical dry forest; PRF, paratropical rain forest; S, successional (probably mostly shrubby); SMDF, successional mesosthermal deciduous forest; SPLDF, successional polar broad-leaved decidous forest; SPRF, successional paratropical rain forest; TRF, tropical rain forest.

‡ D, dry; M, moist, W, wet; S, subhumid; Mg, megathermal; Ms, mesothermal.

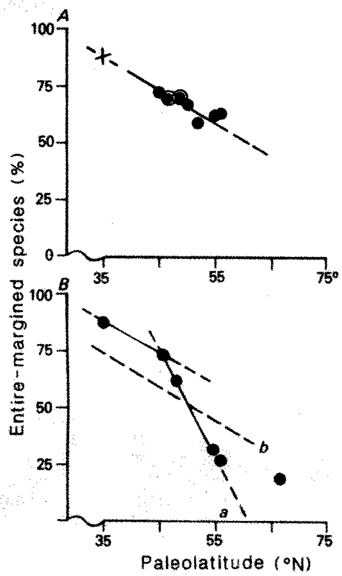


Fig. 3 Least-squares regression analyses of leaf-margin percentages for late Maestrichtian and *B*, early Palaeocene assemblages. In *A*, X denotes early Palaeocene Naborton assemblage and lies along the Maestrichtian regression line. In *B*, line *a* is based on the Raton through eastern Montana assemblages and line *b* is based on all assemblages. Only assemblages of >20 dicot species are included. *A*, $R^2 = 0.55$; *B*, *a*, $R^2 = 0.99$; *b*, $R^2 = 0.24$.

species may be analogous to an extant subgenus, genus or higher-level category[28]. Survival of one species in a previously diverse genus would mask true species-level extinction.

The latitudinal pattern of extinction is that suggested for an 'impact winter'[24]. Most (75%) Raton Lancian leaf species, especially Magnoliidae, are unknown in the early Palaeocene; dispersed cuticles of many of these taxa are found <1 cm below the K-T boundary[16]. Our analyses of the Denver floras indicate similarly high extinction. As at Raton, Denver deciduous taxa had far less extinction than evergreen taxa. In the northern high plains extinction across the K-T boundary was 50-60%[24]. Notable in the early (and later) Palaeocene of the northern High Plains are deciduous taxa that were either uncommon in Lancian mesothermal vegetation or characteristic of polar deciduous forest (for example, trochodendroids and hamamelidaleans). Some evergreens apparently re-invaded from the south. Typically these taxa (for example, *Myrtophyllum*) are the few that occurred in both megathermal and mesothermal vegetation during the Lancian and thus may have survived in the south to re-expand northward temporarily; these taxa are typically rarer in the earliest Palaeocene than in the Lancian and unknown in the late early Palaeocene. Our preliminary analyses of known leaf floras in central Alberta indicate that most (76%) Lancian leaf lineages survived into the early Palaeocene. Extinction was highest among archaic deciduous and evergreen gymnosperms. However, the general pattern in the western interior is extinction of evergreen and survival of deciduous plants. Abilty to enter dormancy was a major factor in survival across the K-T boundary.

In the Raton Basin vegetational change over an estimated 1.5 Myr mimics a successional pattern in diversity (quasisuccession)[16]. This pattern resulted from high extinction in megathermal evergreen vegetation at the K-T boundary. Mesothermal evergreen vegetation also suffered major extinction in the northern high plains, resulting in establishment and persistence of quasisuccessional, low-diversity deciduous forest analogous to successional vegetation in extant mesothermal regions. Polar deciduous forest was disrupted in the earliest Palaeocene but suffered little extinction and was re-established by the late early Palaeocene.

Occupation of the mesothermal region by deciduous forests through most of the Palaeocene[29] indicates that there were no mesothermal refugia; otherwise, evergreens should have rapidly re-occupied this region. Replenishment of evergreens in mesothermal vegetation could then only occur by evolutionary adaptation from an impoverished megathermal flora. Slow replenishment of megathermal vegetation, evidenced by low diversity in Raton Phase 5 assemblages, suggests low survival rate in megathermal refugia.

High survival rate among deciduous versus evergreen plants indicates an event that favoured deciduous plants over evergreens sensitive to low temperature. The simplest explanation is a low-temperature excursion. High survival rate among aquatic ectotherms[30] places major constraints on severity and duration of the temperature excursion. Large aquatic ectotherms might survive a few months of a mean near freezing, but temperatures that would freeze water to a considerable depth, or a long period of inactivity without feeding would certainly be fatal (ref. 31 and J. H. Hutchison, personal communication). Many extant megathermal plants cannot tolerate temperatures <10 °C and most mesothermal broad-leaved evergreens cannot survive a cold-month mean <1 °C (ref. 10). Biological data are consistent with a short period (1-2 months) of a mean temperature near 0 °C. Deciduousness selected by the temperature excursion left a profound imprint on Northern Hemisphere Cenozoic vegetation[29,32]. Rarity of deciduousness in Southern Hemisphere Cenozoic vegetation suggests that the temperature excursion had little effect in the Southern Hemisphere.

The sharpness and magnitude of vegetational and floral changes at the K-T boundary are dramatic. Mass-kill, ecological disruption and climatic change are associated with the event that produced the iridium-rich clay. The clays at various sites must be synchronous: were they of disparate ages, short-term successions, not quasisuccession, would have resulted. Although some[33] have proposed that the event was a marked and prolonged cooling from massive volcanic eruptions, the data from the land biota overall support a brief low-temperature excursion resulting from a bolide impact[1].

We thank W. A. Clemens, A. R. Sweet and officials of the Luscar-Sterco Mine for assistance in various aspects of fieldwork; T. M. Bown, W. A. Clemens, L. J. Hickey, A. H. Knoll, D. J. Nichols, C. B. Officer, P. H. Raven and R. A. Spicer for helpful discussions; and R. A. Stockey for access to the Genesee collections.

Received 11 April; accepted 2 September 1986.

1. Alvarez, W. *et al. Geol. Soc. Am. spec. Pap.* **190,** 305-315 (1982).
2. Alvarez, L. W. *et al. Science* **208,** 1095-1108 (1980).
3. Orth, C. J. *et al. Geol. Soc. Am. spec. Pap.* **190,** 423-433 (1982).
4. Pillmore, C. L. *et al. Science* **223,** 1180-1183 (1984).
5. Tschudy, R. H. *et al. Science* **225,** 1030-1032 (1984).
6. Bohor, B. F. *et al. Science* **224,** 867-869 (1984).
7. Lerbekmo, J. F. *et al. Can. J. Earth Sci.* (in the press).
8. Richards, P. W. *The Tropical Rain Forest* (Cambridge University Press, 1952).
9. Wolfe, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **9,** 27-57 (1971).
10. Wolfe, J. A. *Prof. Pap. U.S. geol. Surv.* **1106,** (1979).
11. Wolfe, J. A. in *Paleobotany, Paleoecology and Evolution* Vol. 2 (ed. Niklas, K. J.) 79-101 (Prager, New York, 1981).
12. Givnish, T. in *Topics in Plant Population Biology* (eds Solbrig, O. T., Jain, S., Johnson, G. B. & Raven, P. H.) 375-407 (Columbia University Press, New York, 1979).
13. Coley, P. D. *Ecol. Monogr.* **53,** 209-233 (1983).
14. Fahn, A. *Plant Anatomy* (Pergamon Press, Oxford, 1967).
15. Dorf, E. *Publs Carnegie Instn* **508** (1942).
16. Wolfe, J. A. & Upchurch, G. R. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
17. Smith, A. G., Hurley, A. M. & Briden, J. C. *Phanerozoic Paleocontinental World Maps* (Cambridge University Press, 1981).
18. Wolfe, J. A. & Upchurch, G. R. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* (in the press).
19. Wolfe, J. A. *Am. geophys. Un. geophys. Monogr.* **32,** 357-375 (1985).
20. Batten, D. J., in *Fossils and Climate* (ed. Brenchley, P.) 127-163 (Wiley, New York, 1984).
21. Upchurch, G. R. & Wolfe, J. A. in *Origins of Angiosperms and Their Biologocial Consequences* (eds Friis, E. M., Chaloner, G. W. & Crane, P. R.) (Cambridge University Press, in the press).
22. Bartels, W. S. *Pap. Univ. Mich. Mus. Paleont.* **24,** 73-79 (1980).
23. Hickey, L. J. *Pap. Univ. Mich. Mus. Paleont.* **24,** 33-49 (1980).
24. Hickey, L. J. *Nature* **292,** 529-531 (1981).
25. Wolfe, J. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **30,** 313-323 (1980).
26. Erickson, J. M. *A. Proc. North Dakota Acad. Sci.* **32,** 79-89 (1978).
27. Boersma, A. in *Catastrophes and Earth History* (eds Berggren., W. A. & Van Counvering, J. A.) 247-277 (Princeton University Press, 1984).
28. Muller, J. *Biol. Rev. Cambridge phil. Soc.* **45,** 417-450 (1970).
29. Brown, R. W. *Prof. Pap. US geol. Surv.* **375,** (1962).

30. Clemens, W. A. *Geol. Soc. Am. spec. Pap.* **190,** 407-413 (1982).
31. Hutchinson, J. H. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **37,** 149-164 (1982).
32. Wolfe, J. A. *Paleobiol.* (in the press).
33. Officer, C. B. & Drake, C. L. *Science* **227,** 1161-1167 (1985).
34. Lerbekmo, J. F., Evans, M. E. & Baadsgard, H. *Nature* **279,** 26-30 (1979).
35. Archibald, J. D. *et al. Geology* **10,** 153-159 (1982).
36. Nichols, D. J. & Ott, H. L. *Palynology* **2,** 94-112 (1978).
37. Hotton, C. *6th Int. Palynolog. Conf. Abstr.,* 66 (1984).
38. Sweet, A. R. & Hills, L. V. *6th Int. Palynol. Conf., Abstr.,* 160 (1984).
39. Middleton, M. D. *thesis,* Univ. Colorado (1983).
40. Knowlton, F. H. *Prof. Pap. US geol. Surv.* **130** (1922).
41. Knowlton, F. H. *Prof. Pap. US geol. Surv.* **155** (1930).
42. Berry, E. W. *Prof. Pap. US geol Surv.* **185-F** (1934).
43. Brown, R. W. *Prof. Pap. US geol. Surv.* **189-I** (1939).
44. Bell, W. A. *Bull. geol. Surv. Can.* **13,** (1949).

Growth and longevity of krill during the first decade of pelagic whaling

A. A. Rosenberg, J. R. Beddington & M. Basson

IIED/IUCN Marine Resources Assessment Group, Imperial College, 48 Prince's Garden, London SW7 1LU, UK

Krill (*Euphausia superba* Dana), the dominant member of the zooplankton community of the Southern Ocean, has been the subject of study for many years. As the main prey species of the baleen whales, several species of seals and many sea bird, fish and cephalopod species, its importance within the ecosystem is clear. It is important for the practicalities of management of the Southern Ocean to estimate the ability of krill to sustain a harvest by man or by its natural predators[1,2]. This estimation requires a knowledge of growth and longevity[3,4], which in krill are the subject of some controversy. Conflicting results have been obtained from krill size composition analysis, which indicates a short lifespan, and age pigment analysis, which indicates a longer life. Here we have analysed eight years of data on krill taken between 1928 and 1938 and have obtained estimates which consistently indicate high growth rates, modulated by a strong seasonal pattern. The cumulative growth through the year is therefore consistent with the studies of age structure using pigment analysis.

In the absence of any method for ageing krill early studies relied on the deduction of estimates of age from information obtained on the length composition of samples[5-11]. The results of these studies indicated a high growth rate, with maximum length being approached within a short life cycle of 2-3 years. Recent work has used the age pigment lipofuscin to determine metabolic age[12,13]. However, these studies suffer from a lack of knowledge of the relationship between metabolic age and real time. They have, to date, only been applied to very small samples of krill. The lipofuscin method, originally developed from studies on insects, has not yet addressed the difficulties of working with an animal like krill which does not have a terminal molt. An unknown proportion of the pigment is lost during ecdysis[12] which may affect the interpretation of results. Despite these problems, these studies indicate a rather longer life cycle with individuals of ages 6-7 being present. The implications for harvesting are that if the shorter lifespan is correct a much larger proportion of krill standing stock may be taken as harvest.

The most extensive time series of length composition samples of Antarctic krill were taken by the British R.R.S. Discovery expeditions which operated from 1926-1951. This sampling period coincided with the first period of major exploitation of the stocks of great whales in the Southern Ocean. We have re-analysed the length composition data on *Euphausia superba* for the period 1928-1938 with two recently developed, simple analytical methods to determine the growth of this species through the period and its expected longevity.

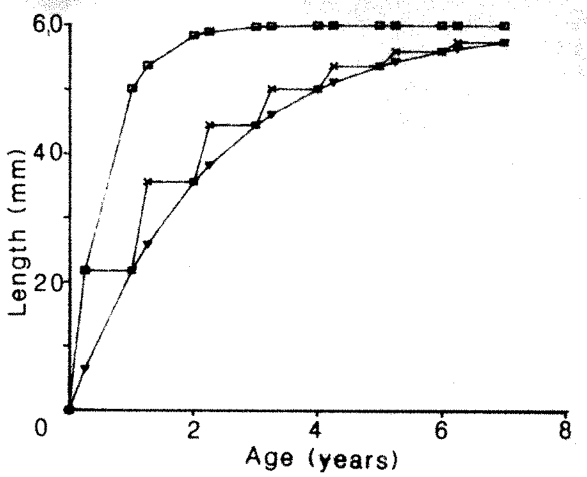


Fig. 1 Von Bertalanffy growth curves with parameters □, $k = 1.8$ (equation 1); ▼, $gk = 0.45$ (equation 2); and ×, the step function for seasonal growth with $k = 1.8$ and $g = 0.25$ (equation 3). In all cases $L_{inf} = 60.0$.

Length composition samples for the period 1928-1938 were taken from the logbooks of the Discovery expeditions and incorporated into a computer database by staff of the Institute of Oceanographic Sciences (Wormley, U.K.) and the Marine Resources Assessment Group of Imperial College (London, U.K.). The population structure of *E. superba* is unknown. Because of this we treated broad geographical regions as the units relevant for analysis. For each year the region containing the most extensive set of samples was used for the estimation of growth rate. We examined the data in two ways: as a standing stock, to provide an average view of the length composition over the year, and to find the monthly change in length composition through the growing season. With a few exceptions the samples were taken between September and May.

We consider growth to be described by the von Bertalanffy curve often used in fisheries biology. That is

$$l_t = L_{inf}\,(1 - e^{-kt}) \qquad (1)$$

where l_t is the length of an animal at age t (where the age of the animal is measured from the time at which it had length zero), L_{inf} is a parameter giving the asymptotic maximum length for the animal and k is a growth constant with units 1/time.

A dominant feature of the Southern Ocean environment is seasonality. There is substantial evidence that growth of many animals virtually ceases during the winter[14,15]. Laboratory experiments on krill indicate that animals may even shrink if starved[16]. To account for such seasonality in samples which average the length composition over the year we modify equation (1) such that

$$l_t = L_{inf}\,(1 - e^{-gkt}) \qquad (2)$$

where g is the proportion of the year in which the animals are growing. We hypothesize that the actual growth of the animal approximately follows a step function given by

$$\begin{aligned} l_t &= L_{inf}\,(1 - e^{-k(t-n(1-g))}) \quad n < t \leq n+g \\ l_t &= l_{n+g} \quad n+g < t < n+1 \end{aligned} \qquad (3)$$

where n is the number of winters the animal has survived. Figure 1 illustrates the growth curves given by equations (1) and (2) and the step function given by equation (3). The monthly length composition data were used to estimate the parameter k in equation (2). The data were the proportion of animals in the total monthly sample in each 5 mm length interval from 5 mm to 65 mm (see Fig. 2, top). To estimate k we write a projection matrix[17], whose elements are a function of the von Bertalanffy equation parameters k and L_{inf}, which takes the length composition at time t to the composition at $t + dt$ where dt is the time between samples. This matrix is analogous to a Leslie population projection matrix. The first sample in a given growth season is projected forward to the sampling times of all subsequent length compositions taken in the same season using the matrix. The

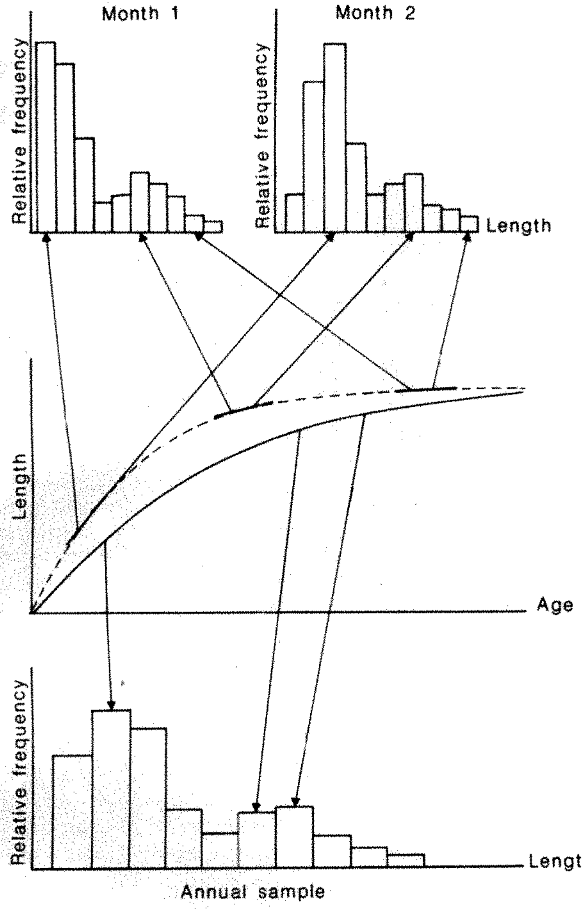


Fig. 2 Schematic representation of growth parameter estimation methods. Dashed line, growth curve for krill during growth season. Arrows indicate movement of individual krill from one population to another. Solid line, length against age for annual sample (age up to 6-7 years). The shape of the dashed line is estimated by examining the length frequency distributions in successive time periods during the growth season, (several months). The parameter *k* which best projects the timecourse of the samples through the season is chosen by least squares. The shape of the solid line is estimated using a method involving a scoring function that determines which parameter combination, *gk*, best match the modes observed in annual length compositions.

estimate of *k* is that value which best matches the predicted to observed length compositions under a least squares criterion (Fig. 2). The value of L_{inf} can be taken as the size of the largest individual observed assuming that the largest individuals in the population are sampled by the gear used. We set L_{inf} to 60 mm for the results reported here. However, our results for the estimation of *k* are insensitive to variations of 5 mm either side of this asssumed value[20].

For the yearly (standing stock) length composition data the parameter estimates are approximately equivalent to the product *gk* in equation (2) since they represent growth averaged over the year, not just within the high growth season. We used a slight technical modification of Shepherd's[18,20] method which scores the modes in the observed length composition based on their predicted position for a given growth curve (Fig. 2). Monte Carlo tests for the two methods indicate their accuracy to be within 10 and 15% of the true value of *k* respectively[20].

The estimated *k* values from the monthly data are quite consistent through the period (Table 1). The average is 1.80. This level of *k* gives a daily growth rate for krill during the fast growth season as 0.148 mm per day. Recent laboratory studies have shown *E. superba* to attain growth rates of around 0.13 mm per day[15,19].

The *k* values for the yearly data (Table 1) again show a high degree of consistency, as with the monthly data. The estimated yearly *k* values are approximately 1/4 of those for the high growth season indicating that maximal growth occurs over approximately three months beginning in the early spring with the seasonal increase in primary production. Such strong seasonality in growth has been reported for other types of animals in Antarctic waters[14]. The strong seasonality in krill growth implies that modes in the length composition for each age group will be most distinct early in the austral spring and samples taken then be most valuable for monitoring growth rate.

Table 1 Growth parameter estimates for krill, *Euphausia superba*, sampled during the Discovery investigations

Split year (June-May)	Summer season growth k*	Summer season growth mm per day†	Yearly growth gk‡	Yearly growth g	Area	n
1928	2.18	0.179	0.43	0.20	Scotia Sea	751
1931	1.79	0.147	0.47	0.26	Scotia-Weddell	994
1934	1.42	0.117	0.27	0.19	Scotia-Weddell	594
1935	1.96	0.161	0.46	0.23	Australian	4,624
1936	1.28	0.105	0.47	0.37	Scotia-Weddell	10,208
1937	2.09	0.172	0.34	0.16	Australian	5,526
1938	1.89	0.155	0.47	0.25	Indian	6,148

Data are for split years over broad geographic regions in the Southern Ocean. Summer season growth refers to the von Bertalanffy parameter *k* estimated from a series of length frequency samples taken in a single growth season. Yearly growth is the estimated parameter combination, *k* times *g*, which represents the average growth over the entire year, not just the growing season. The number (*n*) of krill measured in each split year's samples which were used to estimate the length composition is given.

* Estimated from monthly length compositions—projection matrix.

† At a length of 30 mm growth rate in millimeters per day.

‡ Estimated from yearly length compositions—modal analysis.

Conflicting results concerning the possible lifespan of krill are now interpretable. Growth rates are quite high for this species as would be consistent with a short (2-3 yr) life cycle. However, since the animals only grow through a few months of the year, the effective rate over the entire year is more consistent with a 6-7 yr lifespan as estimated from the age pigment data. Put another way, the average effective growth rate estimated (represented by the step function in Fig. 1) means that it will take a minimum of 6-7 yrs for a krill to approach a length of 60 mm.

An important question is whether the growth of krill has changed in recent years in response to low levels of their baleen whale predators. Our results indicate growth to be quite variable between years and areas, but comparable analyses to our work on the Discovery data have not been performed on recent samples. The recently reported presence of 6-7 year old krill is not inconsistent with length compositions sampled during the first decade of pelagic whaling and so do not imply an increase in krill longevity in response to decreased predation.

This work was supported by the Natural Environment Research Council. A.A.R. was supported by a NATO Postdoctoral Fellowship. We thank Mr A. de C. Baker, Drs I. Everson & J. G. Shepherd for helpful advice and for help in setting up the database. The IIED/IUCN Marine Resources Assessment Group acknowledge support from IUCN and the World Wildlife Fund.

Received 24 April; accepted 10 September 1986.

1. Laws, R. M. *Phil. Trans. R. Soc. Lond.* **B279,** 81-96 (1977).
2. Beddington, J. R. & de la Mare, W. K. in Marine Mammals and Fisheries (eds Beddington, J. R., Beverton, R. J. H. & Lavigne, D. M.) 94-105 (George Allen & Unwin, London, 1985).
3. Gulland, J. A. *F.A.O. Fish. Biol. Tech. Pap.* **97,** 1-425 (1970).
4. Beddington, J. R. & Cooke, J. G. *F.A.O. Fish. Biol. Tech. Pap.* **242,** 1-47 (1983).
5. Ruud, J. T. *Halvråd. Skr. Oslo* **2,** 1-105 (1932).
6. Bargmann, H. E. *Discovery Rep.* **23,** 103-176 (1945).
7. Marr, J. *Discovery Rep.* **32,** 33-464 (1962).
8. Makarov, R. R., Naumov, A. G. & Shevtsov, V. V. in *Antarctic Ecology* Vol. 1 (ed. Holdgate, M. W.) 173-176 (Academic, London, 1970).
9. Ivanov, B. G. *Marine Biol.* **7,** 340-351 (1970).
10. MacKintosh, N. A. *Discovery Rep.* **36,** 1-94 (1972).
11. MacKintosh, N. A. *Discovery Rep.* **36,** 95-156 (1973).
12. Ettershank, G. *Polar Biol.* **2,** 189-193 (1983).
13. Ettershank, G. J. *Crustacean Biol.* **4** (Spec. No. 1), 295-305 (1984).

14. Clark, A. *Mar. Biol. A. Rev.* **21**, 341-453 (1983).
15. Bucholtz, F. in *Antarctic nutrient cycles and food webs* (eds Siegfried, W. R., Condy, P. R. & Laws, R. M.) 339-345 (Springer, Berlin, 1985).
16. Ikeda, T. & Dixon, P. *J. exp. mar. Biol. Ecol.* **62**, 143-151 (1982).
17. Shepherd, J. G. in *The theory and application of length based methods of stock assessment.* (eds Pauly, D. & Morgan, G. P.) Section 1 (ICLARM Conference Series, Manila, in the press).
18. Shepherd, J. G. in *The theory and application of length based methods of stock assessment* (eds Pauly, D. & Morgan, G. P.) Section 2 (ICLARM Conference Series, Manila, in the press).
19. Morris, D. J. & A. Keck *Meeresforschung* **30**, 94-100 (1984).
20. Basson, M., Rosenberg, A. A. & Beddington, J. R. (manuscript in preparation).

Ocular dominance shift in kitten visual cortex caused by imbalance in retinal electrical activity

Barbara Chapman, Michael D. Jacobson, Holger O. Reiter & Michael P. Stryker

Division of Neurosciences and Department of Physiology,
University of California, San Francisco, California 94143, USA

Monocular lid suture during the sensitive period early in the life of a kitten disrupts normal development of inputs from the two eyes to the visual cortex, causing a decrease in the fraction of cortical cells responding to the deprived eye[1]. Such an ocular dominance shift has been assumed to depend on patterned visual experience, because no change in cortical physiology is produced by inequalities between the two eyes in retinal illumination[2] or temporally modulated diffuse light stimulation[3,4]. A higher-level process, involving gating signals from areas outside striate cortex, has been proposed to ensure that sustained changes in synaptic efficacy occur only in response to behaviourally significant visual inputs[5]. To test whether such a process is necessary for ocular dominance plasticity, we treated 4-week-old kittens with visual deprivation and monocular tetrodotoxin (TTX) injections to create an imbalance in the electrical activities of the two retinas in the absence of patterned vision. After 1 week of treatment we determined the ocular dominance distribution of single units in primary visual cortex. In all kittens studied, a significant ocular dominance shift was found. In addition to this physiological change, there was an anatomical change in the lateral geniculate nucleus, where cells were larger in laminae receiving input from the more active eye. Our results indicate that patterned vision is not necessary for visual cortical plasticity, and that an imbalance in spontaneous retinal activity alone can produce a significant ocular dominance shift.

For this study 11 28–34-day-old kittens received intravitreal injections of 5 mM TTX in one eye to block ganglion-cell activity in that eye[6]. Patterned visual input to the other eye was eliminated by suturing the eyelid. Six of the kittens (TTX/lid-suture group) were kept in a normal light cycle, and thus experienced no activity in one eye and diffuse, temporally modulated light in the other. The remaining five kittens (TTX/dark group) were kept in the dark, experiencing no activity in one eye and only spontaneous activity in the other eye. The retinal action potential blockade was maintained for 1 week by repeating the TTX injections. The lid of the injected eye was then sutured and the animals remained or were placed in the dark to prevent vision during recovery of retinal activity. Twenty-four to 48 hours after the retinal blockade began to wear off, the kittens were prepared for electrophysiological recording as described elsewhere[7].

To demonstrate that normal activity had returned to the previously blocked eye, 5–13 single units with central receptive fields (within 15° of the area centralis) were recorded in the lateral geniculate nucleus (LGN) contralateral to the injected eye. In addition, multi-unit responses were assessed every 100 μm through laminae A and A1. No difference in responsiveness was detected between cells driven by the two eyes (see Figs 1, 2), nor between multi-unit recordings in laminae responsive to the two eyes. Briskly responsive ON- and OFF-centre and sustained and transient units were recorded in both laminae.

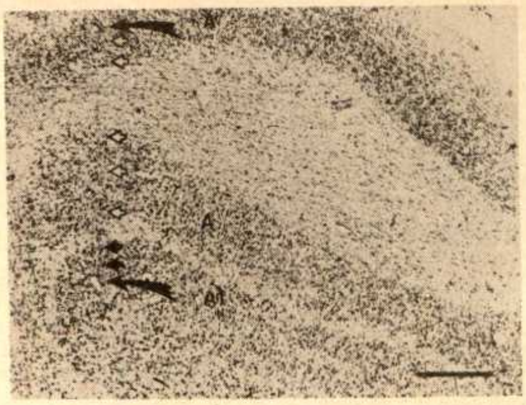


Fig. 1 Coronal Nissl-stained section through an LGN contralateral to the TTX-injected eye. The section shown is 2.88 mm caudal to the rostral pole of the LGN. An electrode track is shown with electrolytic marking lesions indicated by curved arrows. Filled arrows, locations of single units responding to the untreated ipsilateral eye; open arrows, single units responding to the previously blocked contralateral eye. Receptive fields of the units indicated by the top two arrows were at elevation 30°, azimuth 9°; those of the units indicated by the lower five arrows ranged between elevation 6°, azimuth 4° and elevation 4°, azimuth 2°. Scale bar, 0.5 mm.

To determine whether an ocular dominance shift had occurred, we studied cortical area 17 contralateral to the injected eye. This hemisphere was chosen because normal kittens show an ocular dominance bias toward the contralateral eye[6]; thus if recording was in the hemisphere ipsilateral to the TTX-treated eye, an apparent shift in ocular dominance towards the more active eye could be caused, at least in part, by this bias. In the hemisphere contralateral to the injected eye, however, an ocular dominance distribution favouring the more active (ipsilateral) eye must be caused by the effects of deprivation. Horizontal electrode penetrations were used to sample cells across, rather than down, ocular dominance columns; each penetration was long enough to pass through several columns. The intended sample size was 30 cells per animal. Receptive fields of all units studied were within 15° of the area centralis. Ocular dominance was determined for each visually responsive unit using the standard seven-point scale[8]. For both conditions, TTX/lid suture and TTX/dark, ocular dominance histograms were significantly shifted toward the more active, ipsilateral eye (Fig. 3*b*, *c*). These data contrast with histograms from normal kittens of the same age, which show a slight bias in favour of the contralateral eye (Fig. 3*a*), as well as with the more strongly shifted histogram from conventional monocularly lid-sutured kittens (Fig. 3*d*). The degree of shift in each kitten, estimated by a scalar index, indicates that our two groups are significantly different from both normal and conventionally monocularly deprived animals (Mann–Whitney *U* test, $P<0.01$), but that the results in the TTX/lid-suture and TTX/dark groups are not significantly different (Mann–Whitney *U* test, $P>0.05$) (Fig. 4*a*).

In 7 of the original 11 kittens, cross-sectional areas of cells in the LGN were measured in Nissl-stained sections. (LGN cell sizes are known to be smaller in the laminae responding to the deprived eye in kittens with monocular lid suture[1,9] or monocular TTX blockade[10].) As expected, cell sizes were smaller in laminae responding to the TTX-treated eye in both TTX/lid-suture and TTX/dark animals. These morphological results, like the physiological ones, are intermediate between those from normal and monocularly deprived animals (Fig. 4*b*).

Although unlikely, the changes in cortical physiology observed could have been caused by damage to the TTX-treated eye that was not detected in our geniculate recordings, rather

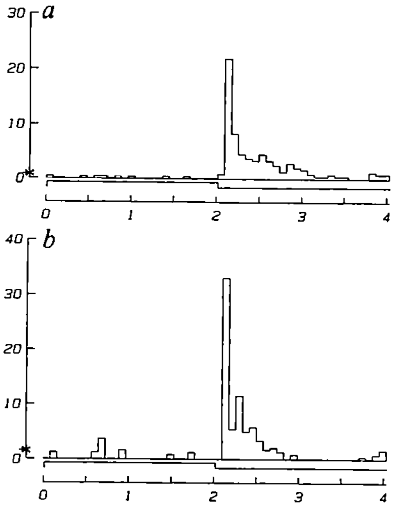


Fig 2 Post-stimulus histograms showing examples of the flash response of single OFF-centre units driven by the untreated ipsilateral eye (*a*) and by the previously blocked contralateral eye (*b*) in the same LGN Stimulus intensity was one log unit above threshold *y*-axis, spikes per s, *x*-axis s Asterisk, level of spontaneous activity, stimulus time course, line above the *x*-axis (ON, up, OFF, down)

than by the activity blockade If such toxicity was responsible for the ocular dominance shifts, the effects should be permanent If, however, differential activity were responsible, then the shifts should be reversible, as are those produced by conventional monocular deprivation in kittens of the same age[11] To test the reversibility of our treatment, two additional kittens were studied These kittens experienced 1 week of the TTX/lid-suture treatment starting at 28 or 32 days old After the TTX blockade wore off, the kittens remained in a normal light cycle for 9 or 11 days with the sutured eyelid remaining closed Thus, these kittens experienced a week of no activity in the TTX-treated eye combined with diffuse light vision through a closed lid in the other eye, followed by a period of conventional monocular deprivation with patterned vision in the previously TTX-treated eye The first week of treatment presumably produced an ocular dominance shift toward the lid-sutured eye, as occurred in TTX/lid-suture animals If TTX injections did not cause permanent damage, this shift should have been reversed during the period of monocular deprivation, resulting in a cortical bias in favour of the eye that had been blocked during the first part of the treatment and had experienced patterned vision during the second part Single-unit recordings from area 17 of the two kittens showed that such a reversal had occurred Ocular dominance histograms from both kittens were strongly shifted toward the eye that was first blocked and then had patterned vision The degree of shift in these kittens (contralateral bias indices 86 and 99, see Fig 4 legend) was comparable to that seen in conventional reverse lid-sutured kittens of the same age[11] Thus TTX injections do not cause lasting damage to the ability of the treated eye to compete for cortical dominance

The similarity of three of the present findings to those from conventional monocular deprivation experiments indicates that the ocular dominance shifts seen in TTX/lid-suture and TTX/dark animals were caused by the differences between activities of the two eyes, rather than by damage to the retinal ganglion cells of the injected eye (1) LGN responses to stimulation of the two eyes were equally good even in the most shifted animals in which few or no cortical cells responded to the injected eye (2) LGN cell sizes were smaller in laminae responding to the injected eye (3) The ocular dominance shifts produced were readily reversible

We draw three conclusions from the results of these experiments First, the significant changes in ocular dominance in both TTX/lid-suture and TTX/dark animals show that patterned vision is not necessary to produce changes in cortical physiology Second, the results from the TTX/dark group indicate that an imbalance in spontaneous activity alone is enough to cause an ocular dominance shift Third, the similar results produced by TTX/dark and TTX/lid-suture treatments suggest that the vision that occurs through a closed eyelid has little additional effect on ocular dominance

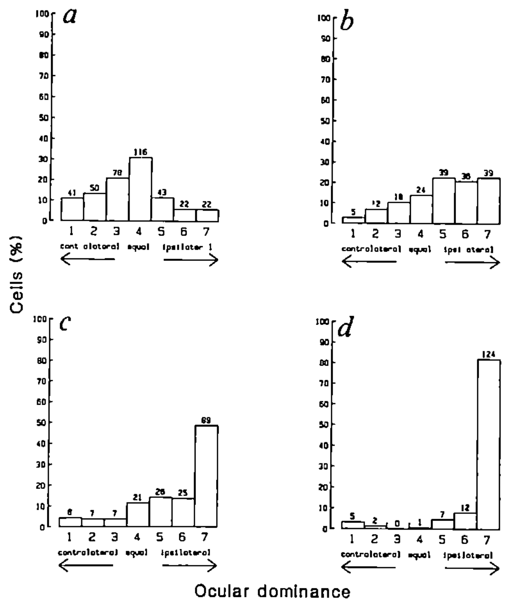


Fig 3 Ocular dominance histograms[8] compiled from single unit responses in area 17 An ocular dominance of 1 indicates a cell driven only by the contralateral eye, 7 a cell driven only by the ipsilateral eye, and 4 a cell driven equally by the two eyes In each deprivation condition, the contralateral eye is less active *a*, 372 visually responsive units from six normal kittens, 36-51 days old 26 visually unresponsive units recorded[6] *b*, 173 visually responsive units from five TTX/dark kittens, 37-44 days old 12 visually unresponsive units recorded *c*, 183 visually responsive units from six TTX/lid-suture kittens, 36-40 days old 19 visually unresponsive units recorded *d*, 151 units from four kittens monocularly deprived by lid suture for 3 5-6 days, followed by 0-2 days recovery in darkness, 29-34 days old[9,15 16]

It is generally believed that cortical plasticity is dependent on cortical cells being driven by their geniculate inputs Silencing all cortical activity through intra-cortical TTX injections prevents ocular dominance shifts in monocularly deprived animals[12] Because geniculate[13] and cortical[14] spontaneous activities are dependent on levels of spontaneous activity in the retina, spontaneous activity must drive cortical cells, however weakly Thus the present results are consistent with a model of plasticity requiring that cortical cells be driven

Our conclusion that ocular dominance shifts can be produced in the absence of patterned vision does not necessarily refute the central gating hypothesis[5] Although shifts in our animals did occur without behaviourally significant visual input to either eye, these shifts were not as large as those seen in conventionally monocularly deprived animals[9,15,16], although the differences in retinal activities were extreme Perhaps the central gate is always slightly open, allowing some degree of shift to occur Behaviourally significant information could further open the gate, causing the larger ocular dominance shifts seen in conventionally monocularly deprived animals We do not know why ocular dominance shifts were produced by the present conditions but not by other rearing conditions with different inputs to the two eyes in the absence of patterned vision[2-4] The actual retinal activities produced by any of these rearing conditions are, however, unknown (except that TTX does block all activity[6]) Therefore, either a large quantitative difference or some temporally modulated qualitative difference between the

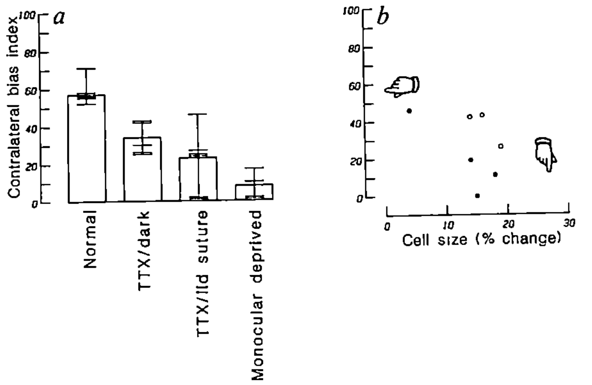


Fig. 4 *a*, Contralateral bias indices showing the range of findings among individual animals from Fig 3 *b*, Scatter plot of per cent change in geniculate cell size versus contralateral bias index *a*, The index provides a single number reflecting the degree of ocular dominance shift for each animal The index is calculated according to the equation

$$100[(1-7)+2/3(2-6)+1/3(3-5)+n]/2n$$

Italicized numbers, number of units in each ocular dominance group, *n*, total number of visually responsive units This index is 100 if all cells are driven exclusively by the contralateral eye and 0 if all cells are driven exclusively by the ipsilateral eye The greater variability among TTX/lid-suture animals is unexplained, it may be caused by differences among the animals of this group in the amount of diffuse-light stimulation of the retina of their sutured eyes *b*, Cells (75 per lamina) were measured in both LGNs of each animal Measurements were made in the latero-medial middle third of coronal sections taken ~2 5 mm caudal to the rostral pole of the nucleus (the area corresponding to the central 5-10° of the visual field)[26] All cell bodies with visible nucleoli were traced in the camera lucida under ×63 or ×100 oil-immersion objectives at final magnifications of ×1,160-1,840 The areas of these tracings were then measured using a Summagraphics Bitpad Planimeter The standard errors ranged from 4 to 6% of the mean for individual laminae Per cent change was calculated as 1−∑ mean cell sizes from laminae responding to the injected eye−∑ mean cell sizes from laminae responding to the more active eye Open circles, data from TTX/lid-suture animals, filled circles, data from TTX/dark animals Upper finger, region of data from normal animals[6], lower finger, region of data from monocularly deprived animals[9] The spread of these control data along the ordinate is shown in *a*, the spread along the abscissa is ~4% for normal animals and 7% for monocularly deprived animals[9] Although there is no strong correlation between the change in cell size and the degree of ocular dominance shift in our animals, the extents of both morphological and physiological changes are intermediate between data from normal and conventional monocularly deprived animals The morphological results for TTX/dark animals (average percentage change = 16) are very similar to those previously reported for older TTX/dark kittens (average percentage change = 17)[10]

activities of the two retinas may be necessary to produce ocular dominance shifts

Our present finding that imbalances in spontaneous activity alone can cause changes in connections to cortical cells, together with earlier work on the effects of spontaneous activity blockade on the development of ocular dominance columns[6,17], is consistent with the hypothesis that spontaneous electrical activity plays a part in the normal fetal development of the visual system[18] Activity-dependent rearrangements of ocular dominance columns, which occur postnatally in the cat[19], begin 3-6 weeks prenatally in the monkey[20,21] If the development of ocular dominance columns follows similar principles in the two species[22], it is not surprising that spontaneous activity plays this part in the cat because in the monkey *in utero*, vision is occluded, and the only possible electrical activity is spontaneous Observations that retinogeniculate axons make functionally effective binocular connections prenatally[23], some of which are lost before birth[23,24], suggest that spontaneous activity is also involved in the refinement of these connections[24,25]

Received 28 January, accepted 23 July 1986

1 Wiesel, T N & Hubel, D H *J Neurophysiol* **26**, 1003-1017 (1963)
2 Blakemore, C *J Physiol, Lond* **261**, 423-444 (1976)
3 Singer, W, Rauschecker, J & Werth, R *Brain Res* **134**, 568-572 (1977)
4 Wilson, J R, Webb, S V & Sherman, S M *Brain Res* **136**, 277-287 (1977)
5 Singer, W in *The Neurosciences Fourth Study Program* (eds Schmitt, F O & Worden, F G) 1093-1110 (MIT, Cambridge, 1979)
6 Stryker, M P & Harris, W *J Neurosci* **6**, 2117-2133 (1986)
7 Schopmann, A & Stryker, M P *Nature* **293**, 574-576 (1981)
8 Hubel, D H & Wiesel, T N *J Physiol, Lond* **160**, 106-154 (1962)
9 Movshon, J A & Dursteler, M R *J Neurophysiol* **40**, 1255-1265 (1977)
10 Kuppermann, B D & Kasamatsu, T *Nature* **306**, 465-468 (1983)
11 Movshon, J A *J Physiol, Lond* **261**, 125-174 (1976)
12 Reiter, H O, Waitzman, D M & Stryker, M P *Soc Neurosci Abstr* **11**, 463 (1985)
13 Eysel, U T & Wolfherd, E *J comp Neurol* **229**, 301-309 (1983)
14 Harris, W A & Stryker, M P *Soc Neurosci Abstr* **3**, 1785 (1977)
15 Hubel, D H & Wiesel, T N *J Physiol, Lond* **206**, 419-436 (1970)
16 Olson, C R & Freeman, R D *J Neurophysiol* **38**, 26-32 (1975)
17 Stryker, M P *Soc Neurosci Abstr* **7**, 842 (1981)
18 Stryker, M P in *Developmental Neurophysiology* (eds Kellaway, P & Purpura, D) (Johns Hopkins University, in the press)
19 LeVay, S & Stryker, M P in *Aspects of Developmental Neurobiology* (ed Ferrendelli, J A) 83-96 (Soc Neurosci, 1979)
20 LeVay, S, Wiesel, T N & Hubel, D H *J comp Neurol* **191**, 1-51 (1980)
21 Rakic, P *Nature* **261**, 467-471 (1976)
22 Des Rosiers, M H *et al Science* **200**, 447-449 (1978)
23 Shatz, C & Kirkwood, P *J Neurosci* **4**, 1378-1397 (1984)
24 Dubin, M W, Stark, L A & Archer, S M *J Neurosci* **6**, 1021-1036 (1986)
25 Sur, M, Garraghty, P E & Stryker, M P *Soc Neurosci Abstr* **11**, 805 (1985)
26 Sanderson, K J *J comp Neurol* **143**, 101-118 (1971)

Intracellular calcium ions decrease the affinity of the GABA receptor

Inoue, M., Oomura, Y., Yakushiji, T. & Akaike, N.*

Department of Physiology, Faculty of Medicine,
Kyushu University 60, Fukuoka 812, Japan

Intracellular free Ca^{2+} ($[Ca^{2+}]_i$) plays a crucial role in the transduction of extracellular signals[1] It has been implicated in the modulation of light sensitivity in *Limulus* photoreceptors[2] and in the efficacy of synaptic transmission[3-5]; calcium ion fluxes are also involved in the postsynaptic facilitation of nicotinic transmission seen in sympathetic ganglia, and in activation of the acetylcholine (ACh) receptor[6,7] $[Ca^{2+}]_i$ is also a second messenger for many biologically active substances[8]. We recorded neuronal activities of sensory neurones from the bullfrog (*Rana catesbiana*), using the suction pipette method[9] and a 'concentration clamp' technique[10] to apply γ-aminobutyric acid (GABA) to the cell. We report the first evidence that $[Ca^{2+}]_i$ suppresses the GABA-activated Cl^- conductance, by decreasing the apparent affinity of the GABA receptor.

Sensory neurones were isolated from dorsal root ganglia by both mechanical dissection and enzymatic digestion in normal Ringer solution containing 0 3% collagenase and 0 05% trypsin at 37 °C for 15 min The isolated cells were maintained in a solution containing equal parts of Ringer and Eagle's minimum essential medium at room temperature (~22 °C) The concentration clamp technique enables complete exchange of the external solution around the cell to occur within a few ms (ref 10) K^+ and Na^+ in both internal and external solutions were replaced with equimolar Cs^+ and Tris, respectively, unless otherwise stated All experiments were performed at room temperature The data have been analysed statistically using Student's *t*-test

Under current clamp conditions, rapid application of GABA (10^{-5} M) in the absence of external Ca^{2+} evoked a large depolarization, and a relatively stable plateau was reached within 2 sec (Fig 1a) This response is mediated through an increase in Cl^- conductance[11] When the application of GABA was preceded by a series of Na^+ spikes (Fig 1b), no effect on the GABA response was apparent However, elicitation of a single Ca^{2+} spike in medium containing Ca^{2+} markedly reduced

---

* To whom correspondence should be addressed

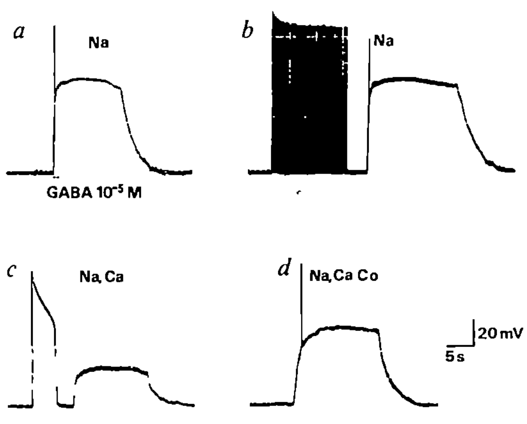


Fig. 1 Effects of Ca^{2+} spike and Na^+ spikes on the GABA-induced depolarization. *a*, 10^{-5} M GABA depolarized the membrane by 75 mV in Na^+ solution (in mM) NaCl 83, $MgCl_2$ 5, TEA-Cl 25, CsCl 2 and glucose 5. The membrane potential was held at −80 mV. *b*, The preceding 50 Na^+ spikes elicited by depolarizing pulses at a frequency of 3.3 Hz had little influence on the 10^{-5} M GABA response (74 mV) in an external Na^+ solution. *c*, Marked suppression of 10^{-5} M GABA response after a 4.3 s duration Ca^{2+} spike in a Na^+ and Ca^{2+} external solution where 2 mM $CaCl_2$ replaced 5 mM $MgCl_2$ and 6 mM NaCl was added to equalize osmolarities, depolarization of 32 mV. *d*, With suppression of the Ca^{2+} spike by adding 4 mM $CoCl_2$ to the Na^+ and Ca^{2+} external solution, 10^{-5} M GABA produced a depolarization of 62 mV which was slightly smaller than the control, probably due to Co^{2+} interference. GABA equilibrium potential of 4 mV agreed well with the theoretical value estimated from the ionic activities of Cl^- in the external and internal solutions. This result suggests that the internal perfusion was complete. The ionic compositions of the standard solutions were (mM) internal, CsCl 95, (TEA-Cl) 25, Cs-aspartate 10, 0.5 EGTA-Ca buffer ($[Ca^{2+}]_i = 3\times10^{-8}$ M) and pH 7.2, external, Tris-Cl 89, TEA-Cl 25, CsCl 2, $CaCl_2$ 2, glucose 5 and pH 7.4. The pH of all solutions was adjusted with Tris and HEPES buffers.

the subsequent response to GABA (Fig. 1c). This inhibitory effect was not simply due to the presence of Ca^{2+} in the external perfusate, since it could be reversed by adding Co^{2+}, which blocks the voltage-dependent Ca^{2+} current (I_{Ca})[12] (Fig. 1d).

Reduction of the GABA response by the Ca^{2+} influx was quantified in Na^+-, K^+-free external solution containing 2 mM Ca^{2+}. Cells were voltage clamped to a holding potential of −50 mV and I_{Ca} was evoked by 50 mV depolarizing command pulses of varying duration. As can be seen in Fig. 2a, the GABA responses were suppressed to a greater extent when the Ca^{2+} influx was increased. A typical relationship between the total amount of I_{Ca} ($\int I_{Ca}$) and the GABA-induced peak response is illustrated in the lower graph of Fig. 2a. It is evident that I_{Cl} in this cell is maximally inhibited when the Ca^{2+} influx exceeds 80 pC. The degree of this I_{Ca}-mediated suppression of I_{Cl} depended on the chelating potency of the internal solution used. When cells were perfused with a solution containing 2.5 mM EGTA, the subsequent response to 10^{-5} M GABA was reduced to 75% ($n = 15$) of control by the preceding 100 ms I_{Ca}, whereas in neurones perfused with 0.5 mM EGTA-Ca^{2+} buffer ($[Ca^{2+}]_i \simeq 3\times10^{-8}$ M), the I_{Cl} was reduced to 50% ($n = 14$) ($p < 0.01$).

An increase of $[Ca^{2+}]_i$ facilitates the onset of depolarization in response to ACh in the frog endplate[13]. In the present study, a similar phenomenon was observed when the time interval between I_{Ca} and the GABA response was short (10 ms as shown in Fig. 2b). In this case, the GABA response reached a maximal value within 250 ms and thereafter I_{Cl} decreased gradually, whereas in the control it continued to rise. However, this acceleration of the GABA-response decay by I_{Ca} was inversely related to the time interval between I_{Ca} and GABA application and was not detectable with an interval of 600 ms, indicating that this facilitated desensitization after an I_{Ca} is only 'apparent'. This phenomenon could be accounted for by the time-dependent activation of the suppressive mechanism triggered by an increase in $[Ca^{2+}]_i$. Figure 2 shows that following I_{Ca}, full inhibition was only achieved after a few hundred milliseconds. The lower graph in Fig. 2b illustrates that recovery from the Ca^{2+}-mediated inhibition occurs within a few tens of seconds. It is not certain whether this time course is one of a change in $[Ca^{2+}]_i$ after I_{Ca} or release of Ca^{2+} from the receptive sites of GABA receptor complexes or a combination of both. In experiments with Ca^{2+} indicator dyes, increases in $[Ca^{2+}]_i$ are reported to last for over 10 s after 1 s I_{Ca} under voltage clamp conditions in sympathetic neurones[14]. These data, however, should not be directly compared with our results, because we recorded neuronal activity with a suction electrode.

It is probable that by dialysis, we removed the soluble cytoplasmic constituents that are important for physiological Ca^{2+} sequestration. However, this disadvantage is not expected to be significant in Ca^{2+} buffering, because in our experiments, cells were perfused with EGTA-containing solutions.

It has been reported that ACh receptor/Cl^- channel complexes in *Aplysia* neurones are inactivated by intracellular Ca^{2+} binding to the receptor complexes[15]. However, Ca^{2+}-mediated inhibition of the Cl^- response in this preparation did not depend on the concentration of ACh, a result which markedly differs from ours. As shown in Fig. 3, the GABA-activated I_{Cl} was inhibited in a competitive manner by I_{Ca}. The degree of this suppression did not depend on membrane potential (which ranged from −40 to −100 mV), although the Ca^{2+}-dependent Cl^- or K^+ conductance[16,17] is markedly voltage-dependent,

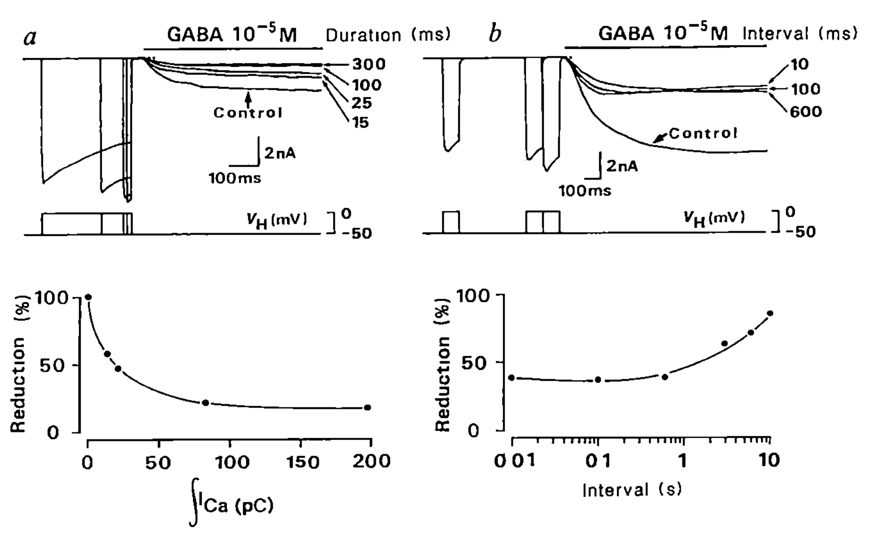


Fig. 2 Suppressing action of I_{Ca} on the GABA-activated I_{Cl}. *a*, Top, actual records of 10^{-5} M GABA-gated I_{Cl}, with and without a preceding I_{Ca} elicited by the depolarizing pulse from the holding potential (V_H) of −50 mV to 0 mV for various durations (15, 25, 100 and 300 ms). The leakage and capacitative currents associated with the ionic currents were subtracted during the experiments by adding the current responses to equal but opposite voltage steps using a signal averager. Bottom, relationship between GABA-gated I_{Cl} and level of voltage-dependent Ca influx (pC). The suppression of GABA-gated I_{Cl} reached a plateau with increasing amounts of I_{Ca}. *b*, Top, actual records of 10^{-5} M GABA-gated I_{Cl} elicited at various intervals after I_{Ca}. Note that the rising phase of I_{Cl} was slow and 'desensitization' became obscure with increasingly large intervals. Bottom, suppression of GABA responses plotted against the interval between the I_{Ca} and GABA responses.

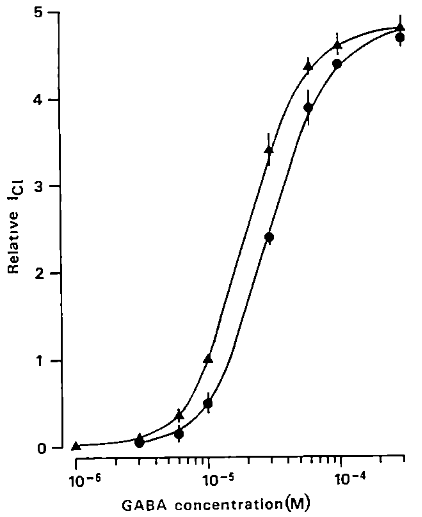


Fig 3 Effect of 100 ms I_{Ca} on the GABA dose-response curve The control dose-response curve (▲) was made by averaging 8 cases, yielding a good fit to the theoretical expression assuming a Hill number of 2 and K_a of 2×10^{-5} M, corresponding to the concentration of GABA producing a half-maximum response When each GABA response was preceded by a 100 ms I_{Ca} (●), more than 5 cases at each concentration were averaged and the data points fell on the line assuming a Hill number of 2 and K_a of 3×10^{-5} M Each data point represents mean ±s e

These results suggest that suppression of the GABA response is related to a decrease in the affinity of the GABA receptor for its agonist We cannot exclude, however, other possibilities such as changes in channel gating kinetics

The experiments reported here clearly demonstrate that a preceding I_{Ca} markedly suppresses the GABA-activated I_{Cl}, and the inhibitory effects are influenced by the EGTA concentration in the internal solution Therefore, it is reasonable to assume that this suppression may be due to an increase in $[Ca^{2+}]_i$ subsequent to the influx of Ca^{2+} However, I_{Ca} could still inhibit the GABA responses in cells perfused with an internal solution containing up to 2 5 mM EGTA These observations suggest that inhibition of the GABA-activated I_{Cl} by Ca^{2+} might occur at a site that EGTA can not gain access to, as in Ca^{2+}-dependent K^+ conductance in smooth muscle[18]

Biochemical studies[17,18] have revealed that various endogenous substances that reduce the binding of GABA are present in the synaptic membranes of the mammalian central nervous system The relationship between Ca^{2+}-mediated inhibition and these substances is currently under investigation The suppression of GABA sensitivity by $[Ca^{2+}]_i$ may be of great importance in phenomena such as epileptogenesis, where the significant Ca^{2+} entry associated with paroxysmal depolarization shifts, may reduce the effectiveness of recurrent inhibition and therefore both prolong burst discharge and enhance the synchronization of firing To our knowledge, this is the first demonstration that $[Ca^{2+}]_i$ rapidly (within a hundred ms) controls the apparent affinity of the GABA receptor

This study was supported in part by grants 59123005 and 59480107 from The Ministry of Education, Science and Culture (N A) We thank M Ohara (Kyushu University) and J Behrends for comments on the manuscript

Received 17 February, accepted 24 September 1986

1 Nishizuka, Y *Nature* **308**, 693-698 (1984)
2 Fein, A & Lisman, J *Science* **187**, 1094-1096 (1975)
3 Lynch, G , Larson, J , Kelso, S , Barrionuevo, G & Schottler, F *Nature* **305**, 719-721 (1983)
4 Ekerot, C F & Oscarsson, O J *J Physiol, Lond* **318**, 207-221 (1981)
5 Kumamoto, E & Kuba, K *Nature* **305**, 345-346 (1983)
6 Dascal, N , Gillo, B & Lass, Y *J Physiol, Lond* **366**, 299-313 (1985)
7 Marty, A , Tan, Y P & Trautmann, A *J Physiol, Lond* **357**, 293-325 (1984)
8 Rasmussen, H & Barrett, P Q *Physiol Rev* **64**, 938-984 (1984)
9 Akaike, N , Hattori, K , Inomata, N & Oomura, Y *J Physiol, Lond* **360**, 367-386 (1985)
10 Akaike, N , Inoue, M & Krishtal, O A *J Physiol, Lond* (in the press)
11 Hattori, K , Akaike, N , Oomura, Y & Kuraoka, S *Am. J Physiol* **246**, C259-C265 (1984)
12 Ishizuka, S , Hattori, K & Akaike, N *J Membrane Biol*/**78**, 19-28 (1984)
13 Miledi, R *Proc R Soc Lond* **B209**, 447-452 (1980)
14 Smith, S J , MacDermott, A B & Weight, F F *Nature* **304**, 350-352 (1983)
15 Chemeris, N K , Kazachenko, V N , Kislov, A N & Kurchikov, A L *J Physiol Lond* **323**, 1-19 (1982)
16 Miledi, R & Parker, I *J Physiol Lond* **357**, 173-183 (1984)
17 Meech , R W & Standen, N B *J Physiol Lond* **249**, 211-239 (1975)
18 Mitra, R. & Morad, M *Science* **229**, 269-272 (1985)
19 Guidotti, A *et al Proc natn Acad Sci U S A* **79**, 6074-6088 (1982)
20 Yoneda, Y & Kuriyama, K *Nature* **285**, 670-673 (1980)

The t(8; 14) chromosomal translocation occurring in B-cell malignancies results from mistakes in V–D–J joining

Frank G. Haluska, Sheldon Finver, Yoshihide Tsujimoto & Carlo M. Croce

The Wistar Institute, 3601 Spruce Street, Philadelphia,
Pennsylvania 19104, USA

The reciprocal chromosome translocation, t(8; 14), involving the heavy chain locus on chromosome 14 (refs 1, 2) and the c-*myc* oncogene on chromosome 8 (refs 3–5) is a characteristic of the B-cell malignancies Burkitt's lymphoma[6] and acute lymphoblastic leukaemia (ALL)[7] We have cloned and sequenced the t(8; 14) breakpoints of an African Burkitt's lymphoma cell line, P3HR–1, and a pre-B cell ALL cell line, 380 (ref 8). In each case the region of chromosome 8 involved has recombined with a J_H region on chromosome 14 The two sites of breakage on chromosome 8 lie within 70 base pairs (bp) of one another. At each joining site, sequences homologous to the signal sequences thought to be recognized by the V–D–J recombinase[9] were identified, as were N regions[9]. In B-cell chronic lymphocytic leukaemias (B-CLL) carrying the t(11; 14) chromosome translocation[10] and in follicular lymphomas carrying the t(14; 18) translocation[11], the V–D–J recombinase is implicated in the mechanism of chromosomal translocations[11]. We speculate that the same enzymatic mechanism is responsible for the t(8; 14) translocations in African Burkitt's lymphoma and pre-B cell ALL.

We have previously described[8] the establishment of a cell line (380) from a patient with a pre-B-ALL (L2 in the French-American-British classification) carrying the t(8, 14) and t(14, 18) translocations Using probes specific to the C_μ region we were able to show that both translocations involve the heavy chain locus, and to clone the breakpoint for the t(8, 14) translocation, which we denoted λ380j9 (ref 12, and Fig 1) To establish which sequences from λ380j9 derive from chromosome 8, we subcloned the 5′ most *Sst*I fragment (p380j9 0 8SS), and used it to probe a Southern blot of DNA from rodent × human somatic cell hybrids containing only chromosomes 8, 14 or the $14q^+$ chromosome from the African Burkitt's lymphoma cell line P3HR-1 (Fig 2*a*) Hybridization occurred only in lanes containing DNA of human chromosome 8, or of the P3HR-1 $14q^+$ chromosome Hence this sequence derives solely from chromosome 8, and this portion of chromosome 8 is translocated to chromosome 14 in P3HR-1

Using this chromosome 8 probe, we screened DNAs from normal and neoplastic tissues in order to detect rearrangements following translocations (Fig 2*b*) Rearrangement was detected in the *Bam*H1 digest of P3HR-1 DNA, as well as the *Hin*dIII digest of Daudi African Burkitt's lymphoma DNA These results indicate that the breakpoints in the 380 and P3HR-1 cell lines lie within the same *Bam*H1 fragment, while the Daudi breakpoint lies on a *Hin*dIII fragment nearby

We attempted to clone the chromosome breakpoints in these cell lines, and to compare them to the 380 breakpoint clone (ref

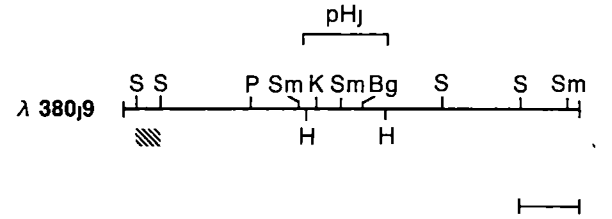


Fig 1 Restriction map of the 14q$^+$ breakpoint clone from the 380 B-ALL cell line, λ380j9 The region of the clone homologous to the joining segment probe pHj (ref 12) is designated by a bracket The hatched box denotes the *Sst*I fragment subcloned as probe p380j9 0 8SS Indicated restriction sites are S, *Sst*I, P, *Pst*I, Sm, *Sma*I, H, *Hin*dIII, K, *Kpn*I and Bg, *Bgl*II Scale bar 2 kb

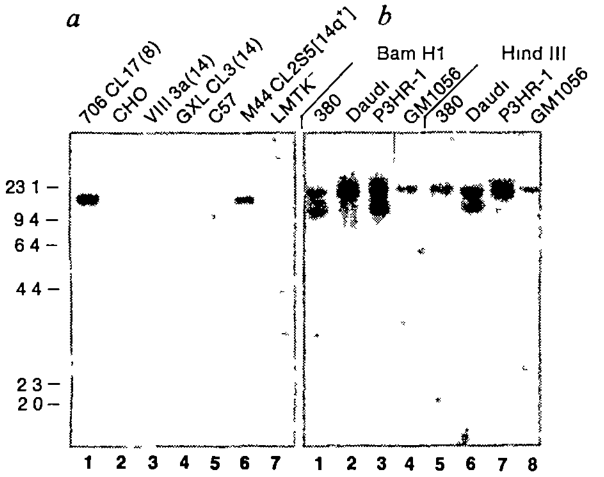


Fig 2 *a*, Southern blot hybridization of *Hin*dIII digested DNAs from human × rodent somatic cell hybrids containing only human chromasome 8, 14 or the P3HR-1 14q$^+$, probed with the p380j9 0 8SS probe Lane 1, 706 Cl 17, a hybrid between a normal human lymphoblast and a Chinese hamster ovary (CHO) fibroblast containing human chromosome 8 but not 14, lane 2, CHO control, lane 3, GM 637×C57 VIII 3a, a hybrid between a simian virus 40-transformed human fibroblast and a mouse parent, containing human chromosome 14 but not 8, lane 4, GM 119×LMTK$^-$ GXL C13, a hybrid between a human fibroblast and mouse LMTK$^-$, also containing human 14 but not 8, lane 5, C56 control, lane 6, P3HR-1×LMTK$^-$ M44 Cl 2S5, a hybrid containing only the 14q$^+$ chromosome from the African Burkitt's lymphoma cell line P3HR-1, lane 7, LMTK$^-$ control *b* Hybridization of 380, Daudi, P3HR-1 and GM 1056 genomic DNAs, with the p380j9 0 8SS probe DNAs were digested with *Bam*HI (lanes 1–4) or *Hin*dIII (lanes 5–8) Lanes 1 and 5, 380 B-ALL cell line, lanes 2 and 6, Daudi Burkitt's lymphoma cell line, lanes 3 and 7, P3HR-1 Burkitt's lymphoma cell line, lanes 4 and 8, GM 1056 lymphoblastoid cell line
Methods Genomic DNA (10 μg) was digested with the indicated restriction enzyme, separated on a 0 7% agarose gel, and transferred to a nitrocellulose filter following denaturation and neutralization The probe was ^{32}P-labelled by nick-translation Hybridization was carried out in 4×SSC, 2×Denhardt's solution and 50% formamide at 37 °C for 24 h, and washing was to a final stringency of 2×SSC at 65 °C Blots were exposed for 12 h at −70 °C with an intensifying screen The size of the λ *Hin*dIII digested marker DNA is in kb

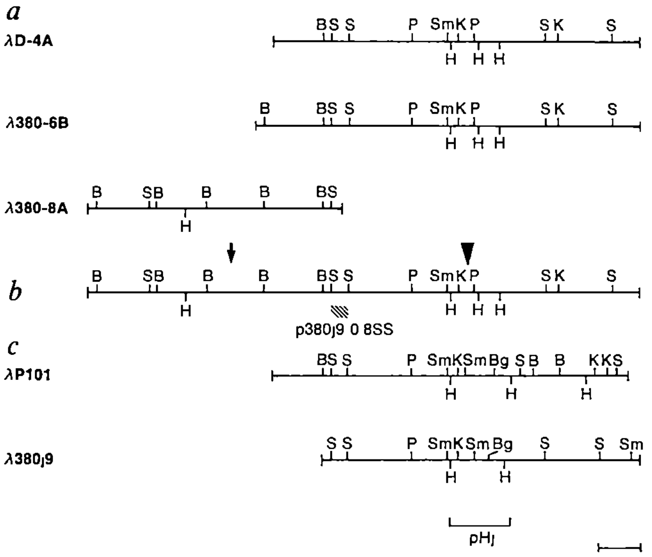


Fig 3 *a*, Restriction maps of the normal chromosome 8 recombinant clones from the Daudi Burkitt's lymphoma (λ D-4A) and the 380 B-ALL (λ 380-6B and 380-8A) cell lines The horizontal lines represent insert DNA from the respective clones *b*, Restriction map of the normal chromosome 8 region deduced from *a* Triangle, P3HR-1 and 380 breakpoints, arrow, putative Daudi breakpoint *c*, Restriction map of the 14q$^+$ breakpoint clone from the P3HR-1 Burkitt's lymphoma (λP101) and 380 (λ380j9) cell lines The point at which the breakpoint clones diverge from the chromosome 8 map is indicated by the triangle in *b* Cleavage sites of the restriction enzymes S, *Sst*I, P, *Pst*I, Sm, *Sma*I, H, *Hin*dIII, K, *Kpn*I and Bg, *Bgl*II Scale bar, 2 kb A representative genomic library was prepared from DNA of P3HR-1 and Daudi cells Cloning procedures in the λ phage EMBL3A (ref 34) were essentially as described[12] Libraries were screened with the p380j9 0 8 SS probe Approximately 350,000 independent recombinant phages in the P3HR-1 library were screened, as well as approximately 800,000 phages in the Daudi library, and 350,000 clones in the previously constructed 380 genomic library[12] The P3HR-1 library was initially screened with the pHj joining segment probe[12] Six recombinant clones were obtained These were secondarily screened with the chromosome 8-specific p380j9 0 8SS DNA probe (Fig 1) Of the six initial positives, one clone, designated λP101, contained homologous sequences

12 and Fig 1) Probe p380j9 0 8SS was used to screen the 380, P3HR-1 and Daudi libraries Two clones were obtained from the 380 library and one each from P3HR-1 and Daudi These clones were subjected to analysis by restriction endonuclease digestion and Southern blotting The resulting physical restriction maps are illustrated in Fig 3 Several features are apparent First, the clones are identical in the region containing the p380j9 0 8SS sequence derived from chromosome 8 Second, the Daudi (λD4) and 380 λ 380-6B and λ380-8A) clones overlap and constitute the normal, uninterrupted chromosome 8 sequence Third, the breakpoints of the 380 (λ380j9) and P3HR-1 (λP101) clones align almost precisely Fourth, although the Daudi breakpoint clone was not obtained, genomic Southern blot data allow us to map the break to the middle of the 3 5 kb *Bam*HI fragment on chromosome 8 (indicated) Finally, comparison of this region with published restriction map[13] of the sequences surrounding c-*myc* on chromosome 8 reveals no similarity ·the breakpoints are situated 5′ to the *Bam*H1 site 15 kb upstream of the first c-*myc* exon

Because the same site on chromosome 8 is involved in the translocation to chromosome 14 in P3HR-1 and 380, we sequenced this region for both the normal and breakpoint clones (Fig 4) The 380 sequence is identical to that of the Daudi normal chromosome 8 allele but diverges six nucleotides 5′ of the *Pst*I site There follows a stretch of 22 nucleotides having homology to neither chromosome 8 nor chromosome 14 sequences The next nucleotide corresponds to the twenty-first nucleotide of the heavy chain joining segment J6 (ref 14), and the remainder of the 380 clone is homologous to this segment, with several single base substitutions The P3HR-1 breakpoint is 63 bp 5′ of the 380 breakpoint on chromosome 8 Upstream of this site, λP101 is homologous to λD4, 3′ it is homologous to a region beginning 90 nucleotides 5′ of J5 This result verifies the somatic cell hybrid data[15] implicating the J_H segments in the P3HR-1 translocation

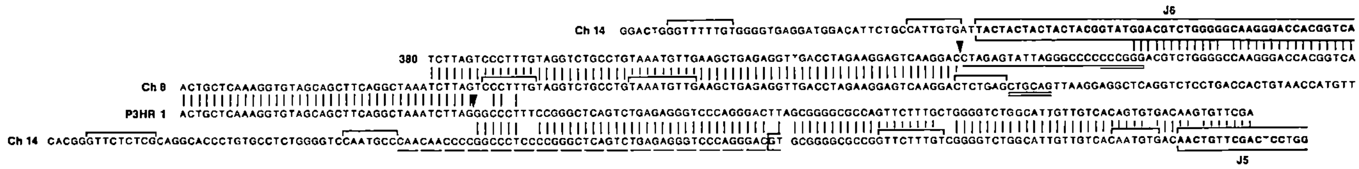

J6
Ch 14 GGACTGGGTTTTTGTGGGGTGAGGATGGACATTCTGCCATTGTGATTACTACTACTACTACGGTATGGACGTCTGGGGGCAAGGGACCACGGTCA
380 TCTTAGTCCCTTTGTAGGTCTGCCTGTAAATGTTGAAGCTGAGAGGTTGACCTAGAAGGAGTCAAGGACCTAGAGTATTAGGGCCCCCCGGGACGTCTGGGGCCAAGGGACCACGGTCA
Ch 8 ACTGCTCAAAGGTGTAGCAGCTTCAGGCTAAATCTTAGTCCCTTTGTAGGTCTGCCTGTAAATGTTGAAGCTGAGAGGTTGACCTAGAAGGAGTCAAGGACTCTGAGCTGCAGTTAAGGAGGCTCAGGTCTCCTGACCACTGTAACCATGTT
P3HR 1 ACTGCTCAAAGGTGTAGCAGCTTCAGGCTAAATCTTAGGGCCCTTTCCGGGCTCAGTCTGAGAGGGTCCCAGGGACTTAGCGGGGCGCCAGTTCTTTGCTGGGGTCTGGCATTGTTGTCACAGTGTGACAAGTGTTCGA
Ch 14 CACGGGTTCTCTCGCAGGCACCCTGTGCCTCTGGGGTCCAATGCCCAACAACCCCGGCCCTCCCCGGGCTCAGTCTGAGAGGGTCCCAGGGACGT GCGGGGCGCCGGTTCTTTGTCGGGGTCTGGCATTGTTGTCACAATGTGACAACTGTTCGAC-CCTGG
J5

Fig. 4 Nucleotide sequence of the joining site between chromosome 8 and 14 in P3HR-1 and 380, and the corresponding normal chromosome 8 region Triangles designate points at which the P3HR-1 and 380 sequences diverge from the chromosome 8 sequence, vertical lines indicate identical nucleotides Bold sequences indicate the J5 and J6 heavy chain joining segments[23] The bracketed regions are conserved heptamer-nonamer signal sequences, and a putative N region is underlined at the 380 joining site The underlined region of chromosome 14 comprises 16 codons and a splice signal (see text) DNA sequences were obtained using the Sanger dideoxy chain termination method[33] on M13[34] subclones from the *Kpn*I site (Fig 3) through the chromosome 14 sequences The *Pst*I site on λD-4A, and the *Sma*I site on λ380j9 (Fig 3) are denoted by double underlines

Both these sequences exhibit features suggesting that the recombinase mediating V-D-J joining was involved in these translocations This enzyme is responsible for the rearrangement of the immunoglobin locus during B-cell maturation[9], and is thought to function similarly in T cells[16] The putative enzyme is thought to recognize heptamer-nonamer signal sequences[17] apposed to each recombining segment Recombination only occurs when a signal sequence having a 12 base spacer is paired with one having a 23 base spacer[18] Moreover, V-D-J recombination results in deletions, substitutions[9], or in the addition of extra nucleotides (N regions), presumably by the action of terminal deoxynucleotide transferase[19] The 380 joining site shows several of these points The breakpoint on chromosome 8 occurs at a sequence homologous to the canonical heptamer (bracketed in Fig 4) Although a satisfactory nonamer is lacking, a heptamer may be sufficient for recombinase function[20,21] The alignment of the J6 segment with the 380 sequence clearly demonstrates that the heptamer-nonamer on chromosome 14 are ideally positioned for a recombination event Additionally, the stretch of extra bases corresponds to an N region, and includes a characteristic stretch of GCs[9] At the P3HR-1 breakpoint, a heptamer is again positioned at the joining site on chromosome 8, twelve bases downstream is a nonamer On chromosome 14, an almost perfect heptamer-nonamer is found, with a 23 bp spacer (shown in Fig 4) Thus recombination according to the 12/23 rule could take place Moreover, the 50 nucleotides 3′ of the nonamer comprise 16 codons, including a stop codon, and a GT splice signal (Fig 4) This sequence has the characteristics of a pseudo-joining segment, and could constitute an additional feature recognized by the recombinase Imprecision is also observed at the breakpoint, as exemplified by the two base mismatch six bases 3′ of the break Lastly, both breakpoints exhibit single base changes which may be a consequence of somatic mutation following V-D-J recombination[9] In both these translocations, then, nucleotide sequence analysis of the sites of chromosome-joining strongly suggests that the translocation occurred in a sequence-specific manner through the action of the recombinase

Recombinase action can be related to the cell of origin of these tumours Both originated in a pre-B cell[22] or sIg^+ cell[20,21] capable of further immunoglobulin gene rearrangement Taken together with similar findings in B-CLL[10] and follicular lymphoma[11], these data suggest that tumour polyclonality[23] should be interpreted as continued gene rearrangement following malignant transformation Moreover, in most sporadic Burkitt's lymphoma cases the chromosome breakpoints are to the first exon of c-*myc*[24-29], and involve the μ, γ, or α switch regions[24-29] However, in most African Burkitt's lymphomas, rearrangements of the c-*myc* locus cannot be detected by Southern blot hybridization[4,30], and translocations into the V_H (ref 2) or J_H (ref 15) region have been found We and others have previously shown that B cells from sporadic Burkitt's lymphomas are more differentiated than those from African Burkitt's lymphomas[30-32] Our results support this hypothesis and suggest that in the African Burkitt's lymphoma P3HR-1, and in the 380 B-ALL, the translocation occurred at the pre-B cell stage of differentiation during heavy chain gene rearrangement The recombinase joined two separated segments of DNA on different chromosomes instead of two segments on the same chromosome Thus in the endemic (African) Burkitt's lymphomas the translocations probably result from mistakes in V-D-J joining while in the sporadic cases it is possible that they result from mistakes during heavy chain class switching in more differentiated B cells

Finally, comparison of the region of chromosome 8 involved in the translocations with the region surrounding c-*myc*[13] shows no overlap Calculations from the extent of the c-*myc* restriction map[13], which covers ~20 kb 5′ of the first exon, and the extent of our clones, which cover at least 15 kb away from the chromosome 14 junction, lead us to conclude that c-*myc* is at least 35 kb distant from the breakpoint in these instances The precedent for transcriptional control acting over such distances comes from the IgH enhancer, which has been demonstrated to act over at least 17 5 kb[33] In each of these translocations, c-*myc* is situated 5′ of the enhancer located between J_H and S_μ This would appear to be the logical explanation for the deregulation of c-*myc* secondary to the t(8, 14) chromosome translocation

We thank Dr Kay Huebner for supplying the somatic cell hybrid DNAs, Dr Masaharu Isobe for many useful comments, Ms Debbie Jiampetti for technical assistance and Mrs K Reinersmann for preparation of this manuscript F G H is a trainee of the Medical Scientist Training Program, NIH 5 T32 GM 07170 Supported by The Outstanding Investigator Award from the National Cancer Institute (C M C)

Received 26 June, accepted 15 September 1986

1 Croce, C M *et al* *Proc natn Acad Sci USA* **76**, 3416-3419 (1979)
2 Erikson, J, Finan, J, Nowell, P C & Croce, C M *Proc natn Acad Sci USA* **79**, 5611-5616 (1982)
3 Dalla Favera, R *et al* *Proc natn Acad Sci USA* **79**, 7824-7827 (1982)
4 Dalla Favera, R, Martinotti, S, Gallo, R C, Erikson, J & Croce, C M *Science* **219**, 963-967 (1983)
5 Taub, R *et al* *Proc natn Acad Sci USA* **79**, 7837-7841 (1982)
6 Manolov, G & Manolova, Y *Nature* **237**, 33-34 (1972)
7 Williams, D L *et al* *Cell* **36**, 101-109 (1984)
8 Pegoraro, L *et al* *Proc natn Acad Sci USA* **81**, 7166-7170 (1984)
9 Tonegawa, S *Nature* **302**, 575-581 (1983)
10 Tsujimoto, Y *et al* *Nature* **315**, 340-343 (1985)
11 Tsujimoto, Y, Gorham, J, Crossman, J, Jaffe, E & Croce, C M *Science* **226**, 1390-1393 (1985)
12 Tsujimoto, Y, Finger, L R, Yunis, J, Nowell, P C & Croce, C M *Science* **226**, 1097-1099 (1984)
13 Bernard, O, Cory, S, Gerondakis, S, Webb, E & Adams, J M *EMBO J* **2**, 2375-2383 (1983)
14 Ravetch, J V, Siebenlist, U, Korsmeyer, S, Waldmann, T & Leder, P *Cell* **27**, 583-591 (1981)
15 Erikson, J, ar-Rushdi, A, Drwinga, H, Nowell, P C & Croce, C M *Proc natn Acad Sci USA* **80**, 820-824 (1983)
16 Yancopoulos, G D *et al* *Cell* **44**, 251-259 (1986)
17 Sakano, H, Huppi, K, Heinrich, G & Tonegawa, S *Nature* **280**, 288-294 (1979)
18 Sakano, H, Kurosawa, Y, Weigert, M & Tonegawa, S *Nature* **290**, 562-565 (1981)
19 Desiderio, S V *et al* *Nature* **311**, 752-755 (1984)
20 Reth, M, Gehrmann, P, Petrac, E & Wiese, P *Nature* **322**, 840-842 (1986)
21 Kleinfeld, R *et al* *Nature* **322**, 843-846 (1986)
22 Korsmeyer, S J *et al* *J clin Invest* **71**, 301-313 (1983)
23 Cleary, M L & Sklar, J *Lancet* i, 1984, 489 (1984)
24 Hamlyn, P H & Rabbitts, T H *Nature* **304**, 135-139 (1983)
25 Gelmann, E P, Psallidopoulos, M C, Papas, T S & Dalla Favera, R *Nature* **306**, 799-803 (1983)
26 Wiman, K G *et al* *Proc natn Acad Sci USA* **81**, 6798-6802 (1984)
27 Hamlyn, P H & Rabbitts, T H *Nature* **304**, 135-139 (1983)
25 Gelmann, E P, Psallidopoulos, M C, Papas, T S & Dalla Favera, R *Nature* **306**, 799-803 (1983)

26 Wiman, K. G *et al. Proc. natn. Acad. Sci. U S A.* **81**, 6798-6802 (1984)
27 Hamlyn, P H & Rabbitts, T H *Nature* **304**, 135-139 (1983)
28 Rabbitts, T H, Hamlyn, P H & Baer, R *Nature* **306**, 760-765 (1983)
29 Showe, L. C *et al. Molec. cell. Biol.* **5**, 501-509 (1985)
30 Pelicci, P G, Knowles, D M, Magrath, I & Dalla Favera, R *Proc. natn. Acad. Sci. U S A.* **83**, 2984-2988 (1986)
31 Benjamin, D *et al J Immun* **129**, 1336-1342 (1982)
32 Croce, C M & Nowell, P C *Adv Immun* **38**, 245-274 (1986)
33 Wang, X H & Calame, K. *Cell* **43**, 659-665 (1985)
34 Murray, N E, Hendrix, R. W, Roberts, J W, Stahl, F, Weinberg, R. A (eds) *Lambda II* (Cold Spring Harbor Laboratory, New York, 1983)
35 Sanger, F, Nicklen, S & Coulson, A E *Proc. natn Acad Sci. U S A.* **74**, 5463-5467 (1977)
36 Messing, J *Meth Enzym.* **101**, 20-79 (1983)

Implications of fragile X expression in normal males for the nature of the mutation

David H. Ledbetter*, Susan A. Ledbetter* & Robert L. Nussbaum†

* Kleberg Cytogenetics Laboratory, Institute for Molecular Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA
† Howard Hughes Medical Institute and Departments of Human Genetics and Pediatrics, University of Pennsylvania School of Medicine and Children's Hospital of Phildelphia, Philadelphia, Pennsylvania 19104, USA

The fragile site at Xq27, associated with a common form of X-linked mental retardation (XLMR)[1], is expressed in a variable proportion of the peripheral lymphocytes of affected males when the cells are cultured under thymidylate stress (Td stress) produced by folate or thymidylate deprivation. Some clinically normal males—transmitting males—are known to carry and transmit the fragile X mutation and yet show no cytogenetic expression in lymphocytes[2,3]. Normal males with no family history of X-linked mental retardation express the site only rarely. When the fragile X chromosome from affected males is isolated in a rodent genetic background by somatic cell hybridization, the level of expression is similar to that seen in lymphocytes under Td stress[4–6]. Here we show that X chromosomes from two transmitting males and two normal control males, all of which were fragile X negative in lymphocytes or lymphoblasts, could be made to express the fragile site in hybrids, although at levels that were below those seen in hybrids from affected males. Furthermore, transmitting males could be differentiated from normal males by their significantly higher expression rates when hybrids were exposed to caffeine[7] before cytogenetic harvest. One male chimpanzee also showed low level expression in hybrid cells. These data suggest that (1) the hybrid system lowers the threshold for fragile X expression, (2) a fragile site at Xq27 may be present on all human and chimpanzee X chromosomes and constitutes a previously unrecognized common fragile site and (3) the hybrid system with caffeine post-treatment can distinguish between the common Xq27 fragile site of control males, the occult mutant fragile site of a transmitting male, and the fully expressed fragile site of an affected male with XLMR. Thus the mutation producing XLMR may represent a multi-step alteration of a naturally occurring DNA sequence producing a continuum of cytogenetic expression and a threshold for clinical manifestation.

We and others have demonstrated that rodent cells (hamster or mouse) are permissive for fragile X expression when somatic cell hybrids are created containing the X chromosome of affected males[4–6,8] This approach allows the isolation of different fragile X mutations on a homogeneous, xenogeneic background Hybrid expression levels are similar to those observed in the human parental cells, suggesting that expression is an intrinsic property of the site and is not influenced by the presence or absence of human autosomal loci[4,6] We have reported preliminary hybrid studies of a clinically normal, transmitting male who showed no expression in lymphocytes or lymphoblasts in multiple laboratories[6,7] Low frequency expression (2%) was observed in hybrids and was enhanced to 12% by caffeine This preliminary observation suggested that the hybrid system lowers the threshold for induction or detection of cytogenetic expression in individuals who carry the mutation but show no cytogenetic expression in routine lymphocyte studies We have analysed additional hybrid clones from this transmitting male and have extended our studies to include a second unrelated transmitting male, two normal control males with no family history of X-linked mental retardation, and one male chimpanzee

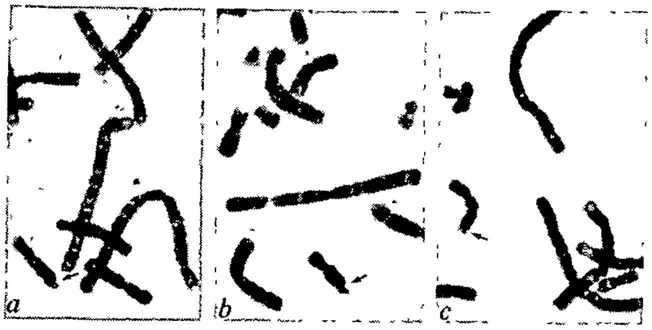


Fig. 1 Partial G-banded metaphases showing fragile X expression in hybrid cells from normal control males N1 (*a*), N2 (*b*), and a chimpanzee (*c*)

One to five hybrid clones from each individual were isolated, characterized, and used for fragile site induction studies The clones studied, parental cell types and their chromosome constitution are shown in Table 1 The number of human chromosomes present varied from 1 (the X only) to 15, with an average of six

Fragile X expression frequencies in parental cells and hybrid cells are shown in Table 2 The two affected males showed lymphoblast expression frequencies of 29% (GM4025) and 40% (DT) over several experiments Under conditions of Td stress alone, hybrid expression was 27% and 23%, respectively These levels were increased to 43% and 30% by caffeine post-treatment Overall the hybrid expression rates were similar to those found in the parental cells

In the two transmitting males, no expression was observed in 200 (MGL66) or 100 (LC) lymphoblast mitoses This is consistent with previous studies, in which lymphocytes and lymphoblasts have been repeatedly negative (refs 9, 10, P A Jacobs, unpublished, W T Brown, unpublished) Under Td stress alone, hybrid cells from each transmitting male showed 2% expression (22/900 and 9/400) over multiple experiments With caffeine, expression increased in the hybrids from both transmitting males to 12% (62/500 and 12/100) This significant enhancement by caffeine is similar to that previously observed in a low-expressing carrier female[6,7]

Several investigators have proposed that the occurrence of clinically normal, transmitting males may be due to the presence of one or more autosomal modifying loci which suppress the major gene on the X chromosome[2,11] Segregation data argue against this hypothesis[12], but no previous experiments have addressed the issue Segregation of the proposed autosomal suppressor by somatic cell hybridization might be expected to produce expression levels equivalent to those of affected males The significantly lower expression rates observed in six independent clones from two unrelated transmitting males suggest an alternative explanation, that the hybrid system lowers the threshold for detection of low level fragile site expression in transmitting males with an occult or incomplete mutation

A total of 100 mitoses from peripheral lymphocytes were scored from each of two normal control males with no family

Table 1 Characterization of somatic cell hybrids used for fragile X expression

Human/chimpanzee parent	Hybrid clone	Hamster parent	Human/chimpanzee chromosomes	Ref*
GM4025	4 12	RJK88	X	4, 6
	PHL3	88 13	X, 3	5
DT	DT10	88 13	X, 3, 4, 7, 10, 16, 17, 20, 21, 22	—
	DT6	88 13	X, 1, 4, 5, 6, 7, 9, 11, 12, 14, 16, 18, 21, 22, ?Y	—
MGL66	MGL66-1	RJK88	X, 13	6
	MGL66-2	RJK88	X, 4, 6, 21, 22, der(9)	—
	MGL66-3	RJK88	X, 4, 11, 12, 15, 20, 21, der(9)	—
	MGL66-4	RJK88	X, 6	6
	MGL66-5	RJK88	X, 12, 14, 15, 18, 21, Y	—
LC	LC1A	RJK88	X, 11, 21, ?Y, ?18	—
N1	RJK739	RJK88	X	—
N2	DHL3A	RJK88	X, 3, 3, 7, 8, 10, 11, 12, 13, 14, 18, 20, 21, 22	—
AG6939A	B4	RJK88	X+approximately 32 autosomes	—

Parental cell types used for fragile X expression and somatic cell fusion were either lymphoblasts, lymphocytes or fibroblasts GM4025 is a lymphoblast line from an affected fragile X male obtained from the Human Genetic Mutant Cell Repository DT is a lymphoblast line produced by Epstein–Barr-virus transformation of an unrelated affected male with a 40% lymphocyte expression rate[1] MGL66 is a lymphoblast line provided by Dr P A Jacobs from a transmitting male who showed no expression in lymphocytes[9] LC is a lymphoblast line provided by Dr W T Brown from a transmitting male with no expression in lymphocytes[10] Peripheral blood was obtained from two clinically normal control males (N1, N2) with no family history of mental retardation for fragile X expression and for fusion A fibroblast cell line on a male chimpanzee (*Pan troglodytes*) was obtained from the Human Genetic Mutant Cell Repository (AG6939A) Hybrids were constructed as previously described[4] from parental cells and RJK88, an HPRT$^-$ Chinese hamster V79 derivative or 88 13, a thymidylate synthase deficient (TS$^-$) mutant of RJK88[5] Hybrids were maintained in Dulbecco's modified Eagle's medium+10% FCS with hypoxanthine (10^{-4} M) and azaserine (5×10^{-4} M) The chromosome composition of each hybrid was determined by trypsin G-banding analysis

* Data are from this work except where stated

Table 2 Per cent fragile X expression in parental and hybrid cells from affected males, transmitting males, normal control males and a chimpanzee male

	Affected		Transmitting		Normal control		Chimpanzee
Cell type/treatment	GM4025	DT	MGL66	LC	N1	N2	AG6939A
Parental—							
Td stress	29 (58/200)	40 (60/150)	0 (0/200)	0 (0/100)	0 (0/200)	0 (0/200)	—
Hybrids							
Td stress	27 (126/468)	23 (34/150)	2 (22/900)	2 (9/400)	2 (5/250)	2 (3/200)	0 2 (1/400)
Caffeine	43 (152/350)	30 (30/100)	12 (62/500)	12 (12/100)	4 (8/200)	5 (9/187)	1 6 (4/250)

The number of fragile X positive cells observed over mitoses scored is given in parentheses Each condition represents at least two experiments For fragile X induction, lymphoblasts and lymphocytes were cultured in medium RPMI 1640+10% FCS and treated with FUdR (5×10^{-8} M and 1×10^{-7} M) or excess thymidine (1-2 mM) for the last 24 hours before standard cytogenetic harvest For hybrids, induction by thymidylate (Td) stress alone was by treatment with FUdR (10^{-9} M-10^{-7} M) or excess thymidine (1-2 mM) in most hybrids, and by thymidine deprivation in the TS$^-$ hybrids[5] for 12-16 hours Caffeine post-treatment was used to enhance expression by addition of 2 mM caffeine for the last 1-2 hours before harvest[7] All slides were trypsin G-banded and at least 50 cells per condition scored for fragile X expression whenever possible Scoring of the lymphocytes and hybrids from the two normal control males was done blindly in the sense that the technicians doing the analysis were unaware of the nature of the experiment and the clinical status of the individuals whose cells they were scoring

history of X-linked mental retardation No fragile X chromosomes were observed However, in blind experiments, 2% expression (5/250 and 3/200) was consistently observed in hybrid cells under Td stress alone This level is identical to that observed in the clinically normal transmitting males Caffeine post-treatment increased expression to 4% (8/200) and 5% (9/187), significantly less than the 12% level of transmitting males ($\chi^2=16\ 8$, $P<0\ 01$) The total of 25 fragile X positive cells observed in the two normal individuals (examples in Fig 1) is much higher than expected by random chromosome breakage Our findings suggest that previous reports of single cell occurrences of fragile X cells in normal controls (reviewed in ref 13) are not artefacts, but represent low level expression of a common fragile site at Xq27 in normal individuals

Common fragile sites are present in both homologues of a chromosome in most or all individuals, whereas rare fragile sites are usually found in just one homologue and occur in less than 1% of the population[14] The relationship between rare and common fragile sites is unclear Their mechanism of induction is similar, though not identical[15] Interestingly, 6 of 18 rare fragile sites occur in chromosome bands which also contain a common fragile site[14] This has led to speculation that a rare fragile site may represent an alteration (mutation) in a pre-existing common fragile site[16,17] Yunis and Soreng[16] demonstrated the presence of common fragile sites in chimpanzee and gorilla in sites homologous to those of man If Xq27 is a common fragile site in humans, it should also be present in these species We therefore extended our hybrid studies to a male chimpanzee No fragile X induction studies were attempted on the parental fibroblast cells, as fragile X expression is generally poor in fibroblasts compared to lymphocytes or lymphoblasts even in affected fragile X males In hybrid cells under Td stress alone, one fragile X cell was observed out of 400 mitoses scored With caffeine post-treatment, 4 positive cells out of 250 were seen (Fig 1*c*) Although this frequency of expression is lower than that seen in normal human males ($\chi^2=4\ 98$, $P<0\ 05$), our impression is that this may be due to the relatively poor growth of this hybrid clone Despite the lower frequency, the presence of a fragile site at Xq27 in a chimpanzee is consistent with the existence of a common fragile site whose origin is ancestral to the human-chimpanzee divergence

Use of caffeine to enhance fragile X expression in hybrid cells was crucial to the discrimination between transmitting and normal males Yunis and Soreng[16] first showed that caffeine, which has synergistic effects with a variety of DNA damaging agents, enhances expression of common fragile sites and fragile X when used in conjunction with fluorodeoxyuridine (FUdR) Although we[7] and others[18-19] have failed to replicate caffeine's enhancing effect on fragile X expression in lymphocytes or lymphoblasts, it has consistently increased expression in hybrid cells[6,7] This difference may be due to the fact that caffeine's effect on fragile X expression appears to be indirect it inhibits the mitotic delay usually associated with DNA damage, thereby reducing the time available for DNA replication, repair and chromosome condensation[7] The effect may vary among cell types with different cell cycle characteristics, particularly different lengths of G_2

The high prevalence and unusual segregation patterns observed in fragile X syndrome suggest that it is not a classical X-linked mutation[3,12,20] The single most unusual characteristic of its inheritance is its variable penetrance within pedigrees For example, the penetrance of mental retardation is significantly lower in offspring of mothers of transmitting males than in offspring of daughters of transmitting males[12] This observation has been referred to as the 'Sherman paradox'[21] and must be explained by any model addressing inheritance of fragile X The unique ability of the hybrid system to induce fragile X expression in transmitting and normal males and to demonstrate a continuum of cytogenetic expression (low in normal males, higher in transmitting males, highest in affected males) argues strongly against the involvement of autosomal modifying loci in this disorder Instead, the presence of a common fragile site at Xq27 is consistent with the multi-step mutation models of Pembrey *et al*[20] and Nussbaum *et al*[22]

We speculate that recombination between two normal X chromosomes at the Xq27 common fragile site region may be the mechanism for an initiation event, producing a clinically normal transmitting male or carrier female Segregation data demonstrates that progression to the full XLMR mutation occurs only in females[20] (who have two X chromosomes), also suggesting that meiotic recombination is involved These recombination events may produce variation in the length or copy number of a DNA sequence at Xq27, which at some threshold level disrupts the function of a neighbouring gene and produces XLMR Quantitative variation in the size of the fragile site DNA region may also explain the variable recombination distances between polymorphic DNA markers observed among fragile X families[23] Thus, our data suggest that a normal DNA sequence at Xq27 may be altered to produce a continuous quantitative variation in fragile site DNA resulting in varying degrees of cytogenetic expression and a threshold for clinical manifestation of the mutation

We thank Drs P A Jacobs and W T Brown for providing lymphoblastoid cell lines on transmitting males This work was supported by NIH RO1 HD20227, the Robert J Kleberg, Jr Center for Human Genetics, and the Howard Hughes Medical Institute R L N is an Associate Investigator in the Howard Hughes Medical Institute

Received 30 May, accepted 26 August 1986

1 Sutherland, G R, Hecht, F, Mulley, J C, Glover, T W & Hecht, B K (eds) *Fragile Sites on Human Chromosomes*, 304 (Oxford University Press, New York, 1985)
2 Froster-Iskenius, U, McGillivray, B C, Dill, F J, Hall, J G & Herbst, D S *Am. J med. Genet* **232**, 619–632 (1986)
3 Howard Peebles, P N & Friedman, J M *Am J hum. Genet* **37**, 956–964 (1985)
4 Nussbaum, R L, Airhart, S D & Ledbetter, D H *Hum. Genet* **64**, 148–150 (1983)
5 Nussbaum, R L, Walmsley, R. M, Lesko, J G, Airhart, S D & Ledbetter, D H *Am. J hum. Genet* **37**, 1192–1205 (1985)
6 Ledbetter, D H, Airhart, S D & Nussbaum, R. L *Am. J med. Genet* **23**, 429–444 (1986)
7 Ledbetter, D H, Airhart, S D & Nussbaum, R L *Am. J med Genet* **23**, 445–456 (1986)
8 Warren, S T & Davidson, R L *Somat Cell molec. Genet* **10**, 409–413 (1984)
9 Rhoads, F A, Oglesby, A C, Mayer, M & Jacobs, P A *Am. J med Genet* **12**, 205–217 (1982)
10 Brown, W T, Gross, A C, Chan, C B & Jenkins, E C *Hum. Genet* **71**, 11–18 (1985)
11 Stembach, P *Hum. Genet* **72**, 248–252 (1986)
12 Sherman, S L, Jacobs, P A, Morton, N E, Froster-Iskenius, U & Howard-Peebles, P N *Hum. Genet* **69**, 289–299 (1985)
13 Jenkins, E C, Brown, W T, Brooks, J, Duncan, C J, Sanz, M M *et al* *Expl Cell Biol.* **54**, 40–48 (1986)
14 Berger, R, Bloomfield, C D & Sutherland G R. *Cytogenet. Cell Genet* **40**, 490–535 (1985)
15 Glover, T W in *Fragile Sites on Human Chromosomes* (eds Sutherland, G R, Hecht, F, Mulley, J C, Glover, T W, Hecht, B K) 80–94 (Oxford University Press, New York, 1985)
16 Yunis, J J & Soreng, A L *Science* **226**, 1199–1204 (1984)
17 Daniel, A *Am. J med. Genet* **23**, 419–427 (1986)
18 Glover, T W, Coyle-Morris, J & Morgan, R *Cancer Genet Cytogenet* **19**, 141–150 (1986)
19 Abruzzo, M A, Pettay, D, Mayer, M & Jacobs, P A *Hum Genet* **73**, 20–22 (1986)
20 Pembrey, M E, Winter, R M & Davies, K E *Am. J med Genet* **21**, 709–718 (1985)
21 Opitz, J M *Am J med. Genet* **23**, 1–10 (1986)
22 Nussbaum, R L, Airhart, S D & Ledbetter, D H *Am. J med. Genet* **23**, 715–722 (1986)
23 Brown, W T, Gross, A C, Chan, C B & Jenkins, E C *Am. J med Genet* **23**, 643–664 (1986)

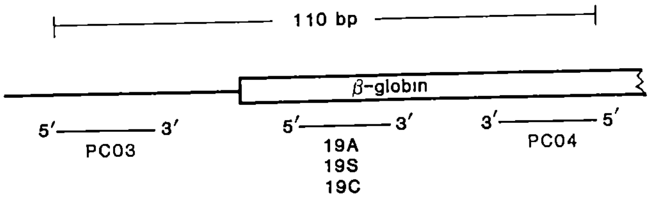
Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes

Randall K. Saiki, Teodorica L. Bugawan,
Glenn T. Horn, Kary B. Mullis & Henry A. Erlich

Department of Human Genetics, Cetus Corporation,
1400 Fifty-Third Street, Emeryville, California 94608, USA

Allelic sequence variation has been analysed by synthetic oligonucleotide hybridization probes which can detect single base substitutions in human genomic DNA[1–3]. An allele-specific oligonucleotide (ASO) will only anneal to sequences that match it perfectly, a single mismatch being sufficient to prevent hybridization under appropriate conditions[4]. To improve the sensitivity, specificity and simplicity of this approach, we used the polymerase chain reaction (PCR) procedure[5] to enzymatically amplify a specific segment of the β-globin or HLA-DQα gene in human genomic DNA before hybridization with ASOs This *in vitro* amplification method, which produces a $>10^5$-fold increase in the amount of target sequence, permits the analysis of allelic variation with as little as 1 ng of genomic DNA and the use of a simple 'dot blot' for probe hybridization. As a further simplification, PCR amplification has been performed directly on crude cell lysates, eliminating the need for DNA purification.

^{32}P-labelled ASO probes have been used in the analysis of a number of genetic diseases caused by single base mutations, including sickle cell anaemia[1], α-1 anti-trypsin deficiency[2] and β-thalassemia[3] Although the ASO method is a powerful diagnostic tool, it requires 5 to 10 μg of genomic DNA, gel-purified high specific activity probes and several days of autoradiographic exposure The short oligomer probes (usually 19 bases long) also hybridize to other genomic sequences so that restriction endonuclease digestion and gel electrophoresis are required to separate the target sequence from the bulk of the genomic DNA In order to develop a simple, sensitive method, we chose the sickle cell anaemia and haemoglobin C mutations in the sixth codon of the β-globin gene[6] as a model system for genetic diagnosis We used ASOs specific for the normal (β^A), sickle cell (β^S), and haemoglobin C (β^C) sequences as probes to detect these alleles in PCR-amplified genomic samples Figure 1 shows the sequence of the PCR primers (PC03 and PC04), the ASO probes (19A, 19S and 19C), and their relationship to the β-globin gene The PCR primers, both 20 bases in length, anneal to opposite strands of the genomic DNA flanking the target sequence[7] These primers are positioned so that the polymerase-catalysed, template-directed extension product of one primer can serve as a template strand for the other In this configuration, repeated cycles of denaturation, primer annealing, and primer extension, result in the exponential accumulation of a 110-base pair (bp) β-globin fragment, with a length defined by the distance between the 5′ ends of the PCR primers For example, 20 cycles of PCR amplification with PC03 and PC04 as primers produce a 220,000-fold increase in the amount of β-globin target



PCR primers

PC03 ACACAACTGTGTTCACTAGC
PC04 CAACTTCATCCACGTTCACC

ASO probes

19C CTCCT$\boxed{A}$AGGAGAAGTCTGC
19A CTCCTGAGGAGAAGTCTGC
19S CTCCTG$\boxed{T}$GGAGAAGTCTGC

Fig 1 Sequence of the PCR primers and ASO probes and their relationship to the β-globin gene. The two PCR primers, PC03 and PC04, are 20 bases long and are complementary to the (−) and (+) strands, respectively. The three ASO probes, 19A, 19S, and 19C, are 19 bases long and span codons 4–17 of the gene[8]. 19A is complementary to the normal β-globin allele (β^A), 19S to the sickle cell allele (β^S), and 19C to the haemoglobin C allele (β^C). Boxes, single base mutations of β^S and β^C.

sequence[7]. The sequences of the 19-base ASO probes used here are identical to those described previously by Studencki *et al*[8].

DNA was extracted from six blood samples from individuals with every possible diploid combination of the β^A, β^S, and β^C alleles and from the cell line GMZ 2064, which has a homozygous deletion of the β-globin gene[9]. The DNA (1 μg from each sample) was subjected to 25 cycles of PCR amplification and 1/30th of the reaction product (33 ng) applied to a nylon filter as a 'dot blot'. Three replicate filters were prepared and each was hybridized with one of the three ^{32}P-labelled ASOs under stringent conditions. The autoradiogram (Fig 2) clearly indicates that each ASO annealed only to those DNA samples which contained at least one copy of the β-globin allele to which the probe was perfectly matched. For example, the β^A-specific probe, 19A, hybridized only to samples AA ($\beta^A\beta^A$), AS ($\beta^A\beta^S$), and AC ($\beta^A\beta^C$). The high degree of specificity contributed by PCR amplification is further emphasized by the lack of any detectable hybridization of the ASO probes to the GM2064 deletion mutant. The frequency of the specific β-globin target in PCR amplified DNA has been estimated by the analysis of cloned amplification products to be approximately 1%, an enrichment of $>10^5$ over unamplified genomic DNA[10]. It is this substantial reduction of complexity that allows the application of 19-base probes in a dot blot format and also makes possible the use of even shorter ASO probes capable of allelic discrimination under less stringent conditions (data not shown).

To estimate the minimum amount of DNA detectable, eight serial dilutions containing 1 to 128 ng of genomic AA DNA were amplified by 25 cycles and half of each sample was applied to a dot blot, then hybridized with the 19A probe. Equivalent dilutions of the 25-cycle amplifications of AA and SS DNA prepared in the previous experiment were included as controls. The autoradiogram (Fig 3) reveals signals in all AA samples, the overall efficiency of amplification with nanogram quantities of template DNA is comparable to that obtained with 1 μg. Moreover, the subnanogram detection limit obtained with ^{32}P-labelled oligonucleotide probes suggests that the analysis of nanogram quantities of genomic DNA with nonradioactive ASO probes (for example, biotin-labelled) should be feasible. Since one nanogram is the amount of genomic DNA present in only 150 diploid cells (assuming a haploid genome size of 3×10^9 bp)

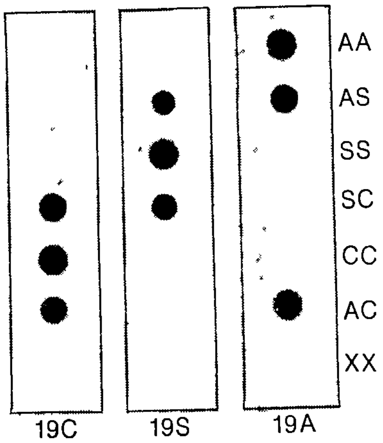


Fig 2 Genotype analysis of PCR-amplified genomic DNA using ASO probes. DNA was extracted from the buffy coat fractions of clinical blood samples of known β-globin genotype[11] and 1 μg portions of AA ($\beta^A\beta^A$), AS($\beta^A\beta^S$), SS ($\beta^S\beta^S$), SC ($\beta^S\beta^S$), CC ($\beta^C\beta^C$), AC ($\beta^A\beta^C$), and GM2064 (deletion mutant, see text) DNA were subjected to 25 cycles of PCR amplification as described previously[7]. Briefly, the samples were dispensed in 100 μl PCR reaction buffer (50 mM NaCl, 10 nM Tris, *p*H 7.6, 10 mM $MgCl_2$, 1 μM PC03, 1 μM PC04, 1.5 mM dATP, 1.5 mM dCTP, 1.5 mM dTTP, and 1.5 mM dGTP). After heating (10 min at 95 °C) to denature the DNA, the samples were transferred to a 30 °C heat block for 2 min to allow the primers to anneal to their target sequences. Primer extension was initiated by addition of one unit of Klenow fragment of *Escherichia coli* polymerase I (New England Biolabs) and incubated for 2 min at 30 °C. This cycle of heating, annealing, and extension was repeated 24 times, subsequent denaturations were by heating for 2 min. The final volume after 25 cycles was 150 μl. One-thirtieth of each reaction, (5 μl, containing 33 ng of the original genomic DNA), was adjusted to 0.4 N NaOH, 25 mM EDTA[17] in a 200 μl volume and applied to a Genatran-45 nylon filter (Plasco) with a Bio-Dot spotting apparatus (Bio-Rad). Three replicate filters were prepared. The ASO probes were phosphorylated at their 5′-termini with [γ-^{32}P]ATP and polynucleotide kinase and purified by spin dialysis[11]. The specific activities of the probes were between 3.5 and 4.5 μCi pmol^{-1}. Each filter was prehybridized individually in 8 ml 5×SSPE, 5× Denhardt, 0.5% SDS for 30 min at 55 °C. Probe (1 pmol) was then added and hybridized at 55 °C for 1 h. Filters rinsed twice in 2× SSPE, 0.1% SDS at room temperature, followed by a high stringency wash in 5×SSPE, 0.1% SDS for 10 min at 55 °C (for 19C) or 60 °C (for 19A and 19S) and autoradiographed for 2.5 h at −80 °C with a single intensification screen (Dupont Lightning Plus).

we considered the possibility that a crude lysate from a small number of cells would contain sufficiently little cellular debris to allow the DNA to serve as template for the PCR amplification procedure. Eight dilutions containing 75 to 9,600 cells from each of the cell lines MOLT-4 ($\beta^A\beta^A$)[11], SC-1 ($\beta^S\beta^S$)[11], and GM2064 were heated at 95 °C to lyse the cells and subjected to 25 cycles of amplification without purification. Equivalent amounts of purified MOLT-4 DNA were included as controls. The amplified samples were divided onto two dot blots and hybridized with the 19A and 19S probes. The autoradiogram (Fig 4) shows a positive signal in an unfractionated cell lysate of only 75 cells. The efficiency of amplification appears to be only two- to three-fold less than with equivalent amounts of purified DNA (MOLT-4 is reported to be 90% tetraploid (A. Green, personal communication) so that 150 cells would be expected to contain 2 ng of genomic DNA). Both ASO probes retain allele specificity. 19A anneals only to MOLT-4 DNA and 19S only to SC-1 DNA. The variation from the expected two-fold reduction of signal intensities may reflect dilution artefacts associated with manipulating such small numbers of cells, and the significant reduction in signal intensity seen in both 9,600-cell samples may be due to inhibition of the amplification process by cellular

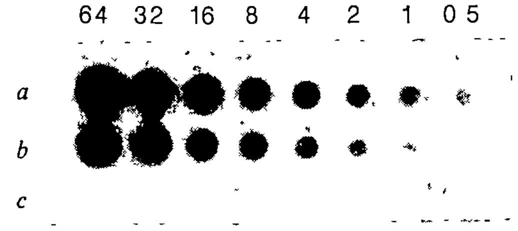


Fig 3 PCR amplification of sub-microgram amounts of genomic DNA *a*, Eight serial dilutions containing 128, 64, 32, 16, 8, 4, 2, and 1 ng of sample AA DNA were amplified by 25 cycles One half of the amplification product (64 to 0 5 ng) was applied to a filter (row A) Controls consisted of equivalent dilutions of *b*, previously amplified AA and *c*, previously amplified SS DNA (Fig 2) The filter was hybridized with the 19A probe and exposed for 6 h See Fig 2 (legend) for methods

debris Use of unfractionated lysate reduces the time required for analysis to <8 hours

PCR amplification and ASO hybridization have also been used to analyse genetic polymorphism at the HLA-DQα locus, which has localized allelic variability in the second exon (Fig 5*a*) Using PCR primers complementary to conserved regions of this exon, a 242-bp fragment of DQα was amplified from DNA samples isolated from nine homozygous typing cell lines (HTC) and from two heterozygous cell lines, then hybridized with four different ASO probes (Figure 5*b*) The sequence of these probes was based on published data[12-15] as well as on our determinations of the sequences of complementary DNA, genomic and PCR-amplified clones (ref 12 and Fig 5, legend) Polymorphism in the HLA-DQ region, which consists of one α and one β chain locus, is detected serologically as the DQw1, DQw2, and DQw3 antigen series[16] The DQA1 probe hybridizes selectively to amplified DNA from all cells typed as DQw1, which is associated with the DR1, DR2, and DRw6 serological types, As expected from the gene sequence, this probe also hybridizes to the DNA from the DR8 homozygous cell line TAB which is also DQw1 The DQA3 probe hybridizes to DNA from the DR3, DQw2 line AVL as well as to the unknown sequence in the DR5, DQw3 line JGL suggesting that these DQα alleles are homologous Since DQA3 did not hybridize to the DNA from the other DQw3 lines (DKB and KOZ), there must be DQα sequence heterogeneity within this specificity Thus the allelic variation in the HLA-DQα locus detected by this method appears to be greater than the polymorphic series defined by serological typing, and can also be more precisely localized This approach is at present being applied to the analysis of polymorphism in the HLA-DQβ and DRβ loci

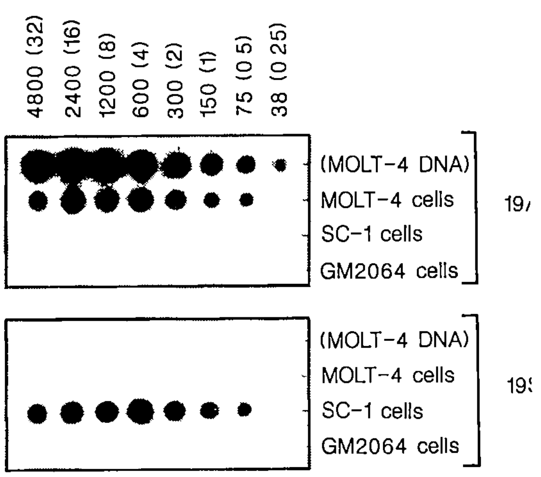


Fig 4 Genotype analysis of PCR-amplified crude cell lysates using ASO probes Cell concentrations of the cell lines MOLT-4 ($\beta^A\beta^A$), SC-1 ($\beta^S\beta^S$), and GM2064 (see text) were determined with a haemocytometer Eight serial dilutions in PBS were prepared at 9 6 × 10^5, 4 8 × 10^5, 2 4 × 10^5, 1 2 × 10^5, 6 0 × 10^4, 3 0 × 10^4, 1 5 × 10^4, and 7 5 × 10^3 cells per ml Ten µl of each dilution (containing 9,600, 4,800, 2,400, 1,200, 600, 300, 150 and 75 cells), were mixed with 35 µl H_2O and heated for 10 min at 95 °C to lyse the cells The lysates were added to 55 µl, 1 8 × PCR buffer (containing salts, primers, and dNTPs) and immediately subjected to 25 cycles of amplification Equivalent amounts (64 to 0 5 ng) of purified MOLT-4 genomic DNA were also amplified The PCR reaction products were divided onto two filters, probed with 19A and 19S and exposed for 6 h Methods are as for Fig 2 except that the amplification buffer included 10% DMSO (v/v), the annealing and extension reactions were performed at 37 °C instead of 30 °C[10] and the PCR amplification procedure was carried out in an automated device designed by the Cetus Instrumentation group

With the use of nonisotopic probes and PCR automation (see Figs 4 and 5), this procedure, combining *in vitro* target

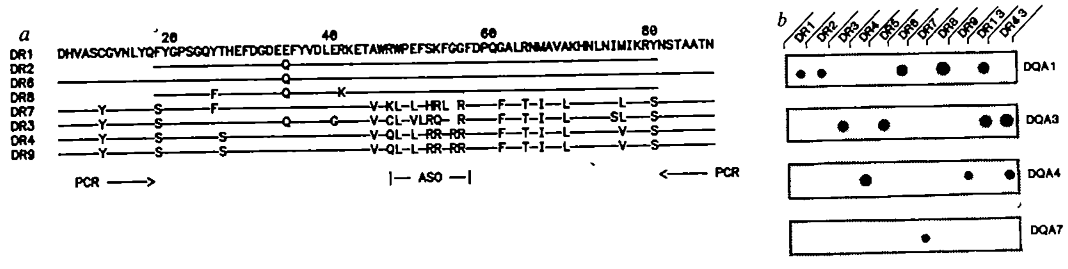


Fig. 5 Detection of allelic variation in HLA-DQα sequences. *a*, Allelic variation in the HLA-DQα gene, *b*, detection of HLA-DQα alleles Protein sequences of the variable outer domain of the HLA-DQα chain, derived from cloned cDNA and PCR-amplified[10] sequences, are shown as differences from the sequence found in the DR1 (ref 13) haplotype (upper line), and grouped to emphasize their similarities The sequences from the DR1, DR2 (ref 13) and DR6 (ref 14) haplotypes (DQw1) are closely related, and those from the DR4 (ref 14) and DR9 (ref 16) haplotypes (DQw3) are identical The region between positions 47 and 56 shows the greatest variation, and was therefore chosen as a target for four different ASO probes In this segment, the DQα sequences from the DR1, 2, 6, and 8 (ref 13) haplotypes are identical and homologous to the DQA1 ASO(5'-AACCTCCAAATTTGCTGAACTCAGGCCACCG-3') The sequence from the DR7 (ref 15) haplotype is homologous to the DQA7 ASO (TTCCACAGACTTAGATTTG) and the DR3 sequence (ref 14) to the DQA3 probe (TGTTTGCCTGTTCTCAGAC) The sequences from the DR4 and DR9 haplotypes are identical and homogous to the DQA4 probe (TTCCGCAGATTTAGAAGAT) *b*, DNA was isolated[11] from nine homozygous typing cell lines and two heterozygous cell lines with HLA-DQα types DR1 (LG2), DR2 (PGF), DR3 (AVL), DR4 (DKB), DR5 (JGL), DR6 (APD), DR7 (LBF), DR8 (TAB), DR9 (KOZ), DR1,3 (B302), DR3,4 (B303) Each DNA sample (1 µg) was PCR-amplified for 28 cycles with the DQα primers as described (Fig 4) The degree of amplification was determined by hybridization to a DQα cDNA probe labelled with ^{32}P by nick-translation, and equal amounts of product (2–10% of the final reaction volume) was denatured with two volumes of 0 5 M NaOH, 1 5 M NaCl After neutralization with 0 5 M Tris, *p*H 7, 1 5 M NaCl to a final volume of 200 µl, the DNA was spotted in four replicates on MSI Nylon 66 (Fischer Scientific) and fixed to the filters by UV irradiation for 4 min The filters were prehybridized for 1 h at 50° in 6 × SSPE, 10 × Denhardt, 0 5% SDS, and hybridized overnight with the ASO probes (at 2 × 10^5 cpm ml^{-1}) labelled with ^{32}P as described (Fig 2), washed for 10 min at 37° in 0 1 × SSPE, 0 1% SDS, and autoradiographed for 90 min with an intensification screen

amplification and ASO probes in a dot-blot format, promises to be a general and simple method for the detection of allelic variation In addition, the ability to analyse genetic variation rapidly in minute amounts of purified DNA or cell lysates has important implications for a wide variety of genetic analyses

We thank Corey Levenson, Lauri Goda and Dragan Spasic for the synthesis of oligonucleotides, Norm Arnheim for discussions, and Larry Johnson and Richard Leath for instrument and engineering support We also thank John Sninsky for critical review, and Kathy Levenson, Eric Ladner, Tim Culp and Sharon Nilson for preparation of this manuscript

Note added in proof Recently, 50 ng samples of PCR-amplified genomic DNA have been analysed using a biotin-labelled ASO probe

Received 30 May, accepted 5 September 1986

1 Conner, B J *et al. Proc. natn. Acad Sci USA.* **80,** 278-272 (1983)

2 Kidd, V J, Wallace, R B, Itakura, K & Woo, S L C *Nature* **304,** 230-234 (1983)

3 Pirastu, M *et al. New Engl. J Med* **309,** 284-287 (1983)

4 Wallace, R B *et al Nucleic Acids Res* **6,** 3543-3557 (1979)

5 Faloona, F A & Mullis, K B *Meth. Enzym* (in the press)

6 Kan, Y W & Dozy, A M *Science* **209,** 388-391 (1980)

7 Saiki, R K *et al Science* **230,** 1350-1354 (1985)

8 Studencki, A B *et al Am J hum Genet* **37,** 42-51 (1985)

9 Tuan, D *et al Proc. natn. Acad Sci USA.* **80,** 6937-6941 (1983)

10 Scharf, S, Horn, G & Erlich, H A *Science* **223,** 1076-1078 (1986)

11 Saiki, R K, Arnheim, N & Erlich, H A *Biotechnology* **3,** 1008-1012 (1985)

12 Bugawan, T, Horn, G, Long, C & Erlich, H A (Manuscript in preparation)

13 Horn, G *et al.* (manuscript in preparation)

14 Chang, H, Moriuchi, T & Silver, J *Nature* **305,** 813-815 (1983)

15 Moriuchi, J, Moriuchi, T & Silver, J *Proc natn Acad Sci USA* **82,** 3420-3424 (1985)

16 Albert, E D, Baur, M P & Mayr, W R. *Histocompatibility Testing,* 4-8 (Springer, Berlin, 1984)

17 Reed, K C & Mann, D A *Nucleic Acids Res* **13,** 7207-7221 (1985)

Common evolutionary origin of legume and non-legume plant haemoglobins

Jorg Landsmann, Elizabeth S. Dennis,
Thomas J. V. Higgins, Cyril A. Appleby,
Alexander A. Kortt* & W. James Peacock

Commonwealth Scientific and Industrial Research Organisation,
Division of Plant Industry, GPO Box 1600, Canberra,
ACT 2601, Australia
* Division of Protein Chemistry, CSIRO, Parkville,
Victoria 3052, Australia

The detection of haemoglobins in distantly-related non-legume plant families[1,2] as well as in the legume families raises the question of whether the origin of the plant haemoglobins was single or multiple. We have isolated a haemoglobin gene from *Parasponia andersonii* (Ulmaceae) which shows more than 50% nucleotide sequence homology with the haemoglobin genes of legume plants and has three introns at identical positions to the leghaemoglobin introns[3-5]. The *Parasponia* gene has homology to haemoglobin genes in another distantly-related nodulating plant, *Casuarina*. It also hybridizes at high stringency to sequences in a related, but non-nodulating genus, *Trema*. We conclude that the globin gene family may be widespread in modern plants, that plant haemoglobins may have a cryptic function in non-symbiotic tissue and that plant haemoglobins have evolved by vertical descent, probably from an ancestor common to modern plants and animals.

A number of plants outside the legume families are known to form nitrogen-fixing root nodules in symbiotic association with *Rhizobium*[6] or actinomycetes[7] Haemoglobin has been detected in nodules of several non-legume plants[1,2] and characterized from *Parasponia* (Ulmaceae)[8,9] and *Casuarina* (Casuarinaceae)[10] Modern systems of plant classification show the known nodulating, haemoglobin-containing species to be widely distributed over several superorders of dicotyledons[7,11,12] (Fig 1) No continuous evolutionary pattern is evident

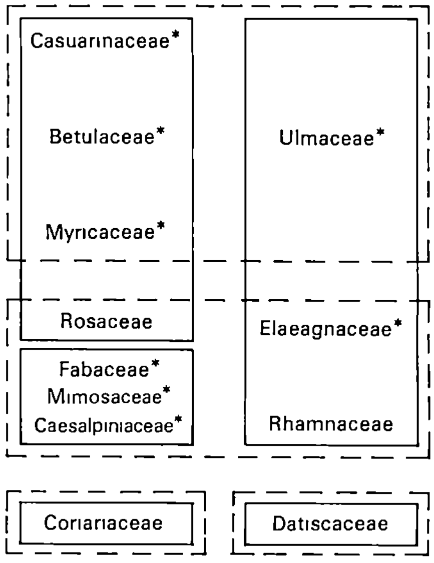


Fig 1 Phylogenetic relationship of dicotyledonous plant families in which some species have been shown to form nitrogen-fixing nodules[7] *Parasponia* is the only genus of the Ulmaceae which has been shown to nodulate, whereas in the legume families nearly all genera of the Fabaceae and Mimosaceae nodulate, the Caesalpiniaceae are mostly non-nodulating Families in which at least one species has been shown to contain nodule haemoglobin[1,2] are marked with an asterisk Solid line, superorders according to Dahlgren[12], dotted line, according to Cronquist[11] Note that these superorders also contain ~150-250 (depending on the classification system) non-nodulating families not listed here The N_2-fixing symbiosis in legumes is always with *Rhizobium* bacteria, *Parasponia* is the only known non-legume genus to be nodulated by *Rhizobium* strains, some of which also infect legumes[6] Other non-legumes undergo symbiosis with *Frankia*, an actinomycete[7]

Although there is 40% protein homology between soybean and *Parasponia* haemoglobin[9], a soybean leghaemoglobin complementary DNA (cDNA) clone did not hybridize to *Parasponia* DNA or nodule RNA, even at reduced stringency A 17mer and a 14mer oligonucleotide sequence deduced from the *P andersonii* haemoglobin I protein sequence[9] were made (see Fig 4) Both oligonucleotides hybridized to an RNA 750 bases long in Northern blots of *Parasponia* nodules (Fig 2*a*) The 17mer was used as primer for synthesis of radioactive cDNA, which was used as a probe, it also hybridized to a 750-base RNA in Northern blots (Fig 2*b*) and identified haemoglobin-specific clones in a nodule cDNA library Two positive clones were analysed, pL209 contained the entire coding region for *Parasponia* haemoglobin I whereas pL201 lacked the first 16 codons Using pL209 as a probe, three independent genomic clones were isolated from a *Bam*HI-EMBL4 genomic library All three clones contained one complete *Parasponia* haemoglobin gene on a 4 8 kilobase (kb) *Eco*RI end fragment within a 13-kb *Bam*HI fragment The sequences of the genomic clones, pL305 and pL401, correspond to the cDNA clones pL209 and pL201 respectively (Fig 3)

The coding region of the *Parasponia* haemoglobin gene spans 486 nucleotides which translate into 161 amino acids excluding the N-terminal Met (Fig 4) Predictions from the DNA sequence showed discrepancies with the published protein sequence[9] Resequencing of the relevant peptides confirmed an additional Phe at position 36, an additional Val at position 73 and a Cys at position 78 instead of the previously reported Thr Four additional amino acids (Ser-Ser-Ser-Glu) were found at the N-terminus (Fig 4) *Parasponia* haemoglobin has the widely conserved globin residues C2-Pro, CD1-Phe, proximal His (E7) and distal His (F8) and, as in other plant haemoglobins, an extended EF-region[13] The protein coding DNA region of the *Parasponia* haemoglobin gene shows more than 50% homology

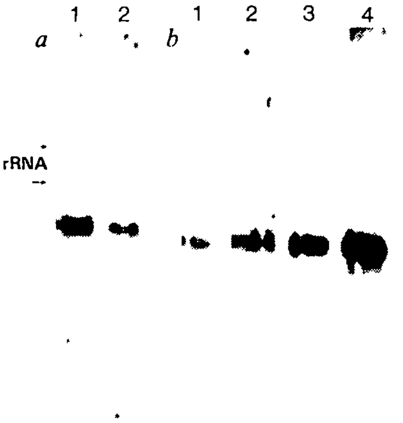


Fig 2 Northern hybridization of *P andersonii* nodule poly(A)$^+$ RNA *a*, Probed with lane 1, the 17mer oligonucleotide, lane 2, the 14mer oligonucleotide *b*, Probed with lane 1, a mixture of 14mer and 17mer, lane 2, 17mer primed cDNA, lane 3, pL201, lane 4, pL209 (see Fig 4 for sequences of the oligonucleotides)
Methods Aseptically-germinated seeds of *P andersonii*, inoculated and nodulated with *Rhizobium* CP283, were transplanted and grown in a sand/vermiculite mixture and cultured using a nitrogen-deficient nutrient medium[6] Root nodules were harvested from a number of plants and RNA was prepared from these according to standard procedures[22] In Northern hybridizations 10 μg of poly(A)$^+$ RNA was run on a formaldehyde agarose gel and hybridized in 50% formamide, 3 5×SSC, 0 75×Denhardt's, 15 mM NaPO$_4$ buffer *p*H 6 5 according to standard procedures[22] Hybridizations with cDNA clones and 17mer-primed cDNA as probes were at 42 °C, but at 32 °C when oligonucleotides were used as probes In oligonucleotide-primed cDNA synthesis, equimolar amounts of poly(A)$^+$ RNA and 17mer were annealed for 30 min at 49 °C in 50 mM Tris/HCl *p*H 8 3, 50 mM KCl, 8 mM MgCl$_2$ The first strand of the cDNA was labelled with [α-^{32}P]dCTP using AMV-reverse transcriptase Oligonucleotides were end labelled with [γ-^{32}P]ATP and T4 polynucleotide kinase

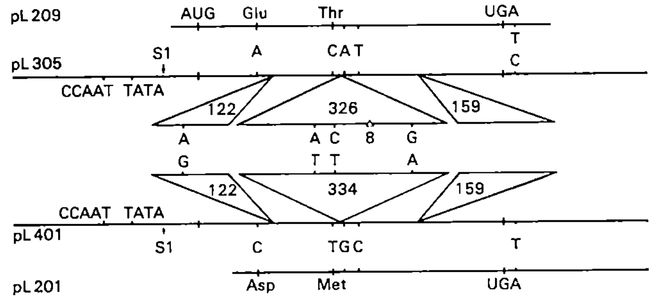


Fig 3 Structure of cDNA clones pL209 and pL201 and genomic clones pL305 and pL401 The cDNA clones contain *Pst*I-*Pst*I inserts of 664 bp (pL201) and 613 bp (pL209) The genomic clones contain a 4 8-kb *Eco*RI fragment each The DNA sequence of pL209 corresponds to pL305, the sequence of pL201 to pL401 Differing nucleotide positions and deduced amino-acid residues are indicated (see also Fig 4) Translational start (ATG) and stop (TGA) codons and presumptive promoter signals for transcription (CCAAT and TATA) and S$_1$-mapped start of transcription are indicated Intron 1 differs by 1 bp between the two genomic clones and intron 2 by 3 bp and an 8-bp deletion/insertion
Methods For cDNA cloning we followed the method of Gubler and Hoffman[23], using oligo(dT) as primer After the final ethanol precipitation in 2 M NH$_4$-acetate the unfractionated double-stranded cDNA population was C-tailed with ^{3}H-dCTP and terminal deoxynucleotidyl transferase, heated for 5 min at 65 °C and annealed for 4 h at 42° to G-tailed pUC8 containing about 12 G residues at *Pst*I ends (gift of W Gerlach) according to standard procedures[22] Transformation of *Escherichia coli* JM83[24] was according to Hanahan[25] Approximately 10^4 white colonies per μg of input *Parasponia* poly(A)$^+$ RNA were isolated on X-gal plates containing 100 μg ml^{-1} ampicillin Colonies were hybridized as described in Fig 3 but with 5×SSC at 42°C with 17mer-primed cDNA as probe Two positively hybridizing clones, pL201 and pL209, were analysed The clones do not contain a poly(A) tail, probably due to the use of RNase H in cDNA cloning A genomic library in λ EMBL4 was made from a 10-15-kb *Bam*HI-fraction of *P andersonii* total DNA This DNA was made from very young leaves taken from a number of plants Plaques were hybridized with pL209 as probe Three independent positive clones (out of 120,000) were identified All contained a 13-kb *Bam*HI fragment containing a 4 8-kb haemoglobin-specific *Eco*RI-fragment These fragments were subcloned into pUC19 from two of the genomic clones giving rise to pL305 and pL401

with leghaemoglobin genes from soybean and kidney bean (Fig 4) DNA homology is higher than the protein homology because amino-acid substitutions in various parts of the haemoglobin protein (for example, positions 18-34, 53-62, 71-78 and 91-100) have, in these regions, allowed mutations in the first codon positions to occur as frequently as in the second and third positions, thus altering the protein faster than its encoding DNA However, the non-coding regions of the *Parasponia* haemoglobin gene have diverged from those of the leghaemoglobin genes Only short sequences, which may be involved in regulatory functions, remain similar

The two genomic *Parasponia* haemoglobin DNA sequences contain three introns (Figs 3 and 4) which are at the same positions as those in the leghaemoglobin genes Introns 2 and 3 are inserted between codons, intron 1 splits a codon after its second base in *Parasponia*, as in the leghaemoglobins[3-5] and in the vertebrate globins[14,15] The central intron is unique to the plant haemoglobins[3-5,14,16] The conservation of these intron positions in legumes and a non-legume plant indicates a common origin for plant haemoglobins from an ancestral functional gene unit having three introns Gō suggested[17], on the basis of protein domain analysis, that a third central intron would also have been present in the putative ancestor of animal globins

In contrast with the haemoglobin multigene families found in legume plants[3], Southern blot analyses of *Parasponia* genomic DNA indicate there is only one haemoglobin gene locus A single band appears in digests with *Bam*HI (13 kb), *Eco*RI (7 kb), *Hin*dIII (9 kb) and *Bgl* II (8 5 kb) with pL209 as probe However, two different haemoglobin clones were isolated from both cDNA and genomic DNA The clones differ at four nucleotide positions in the coding region, two of which led to an amino-acid change, and by a few bases in introns 1 and 2 (see Fig 3) Intron 1 has 122 nucleotides in both genomic clones, intron 2 has 326 or 334 nucleotides and intron 3 is 159 nucleotides long With the exception of a single base the clones are identical over at least 300 base pairs (bp) upstream and downstream in the non-translated sequences These data suggest the existence of allelic heterogeneity at a single locus rather than different genes The fact that *Parasponia* has not developed a multigene haemoglobin family as in the legumes, and that *Parasponia* haemoglobin has an unusually long polypeptide chain may reflect a relatively primitive haemoglobin gene condition[18]

Legume and *Parasponia* haemoglobins both have around 15% protein homology with animal globins and introns 1 and 3 in plant haemoglobins are at positions identical to the introns in vertebrate globins A monophyletic origin of all plant and animal haemoglobins and their vertical evolution seems probable The invocation of horizontal gene transfer[3] between animals and plants appears to be unnecessary but if it occurred it is likely that it predates the diversification of angiosperm plants[19] A presumptive common ancestor of animal and plant haemoglobins could date back 1,500 million years, before the time of the radiation of animals and plants[20]

The *Parasponia* haemoglobin cDNA probe hybridizes to haemoglobin genes in the distantly-related *Casuarina*, which

```
     5'                                                                                              100
Para TTATCTTACTAAAAAGAAAACGAAAATAAAAAACCCAAAGATATGGCTCCCCAATACCCTGAAGAGTTACACACGATCCCCATTTTTTCTACTATATATA

                                         SI                   <pL209                                  200
Para CAGAGTGCCTTCACCAGATTTTCCAAACACACTCCAACATATCCCATTGCCCAAATAAAAATTTCTCAGCTTTTAGTCCCCTCAACCCACAGAAGCCATG
Kid                                                                                                  *GAA

                 HpaI                                 <pL201                                   C      300
Para AGCAGCTCAGAAGTTAACAAAGTTTTCACAGAGGAGCAGGAAGCTCTGGTGGTGAAAGCATGGGCTGTAATGAAGAAGAACTCTGCTGAACTGGGTCTTC
Kid   AGTAGAA AG AAA TGGGT C      T   A    A  G         AAC GCT      AA C T C   GGA   AT C AC  TACA  G  G
                                                                                           Asp
     SerSerSerGluValAsnLysValPheThrGluGluGlnGluAlaLeuValValLysAlaTrpAlaValMetLysLysAsnSerAlaGluLeuGlyLeuG

                                                                   G                                  400
Para AATTCTTCCTCAAGTAAGTCAAAATTATATATAGTACACTTTTTATTTACTTTGCTTCTTTTATAGACCAAGTTTTTGAATAAAAGGGTACTATTTTTTT
Kid  TG    A AC TC<--------------------------------------------------------------------------------------

     lnPhePheLeuLy

                                                                                                      500
Para TTCCTGAAAAAAATTGGTTGATTGAAACTTTGCAGGATATTTGAGATTGCACCGTCTGCCAAGAACTTGTTCTCTTATTTGAAGGACTCTCCGGTTCCTT
Kid  ---------------------------------->    C G    AA     AG A  A     T        A T C TGCTA TGGAGTA A   C+
                                                                           50
                                        sIlePheGluIleAlaProSerAlaLysAsnLeuPheSerTyrLeuLysAspSerProValProL

                                      T                                                               600
Para TGGAGCAGAACCCAAAGCTCAAGCCCCATGCTACGACTGTCTTCGTTATGGTAAAGCCAACTTTTGTTCTCCTATTCCCTTATCCTAATTTTACAAGAAT
Kid  +++++ACT  T  T       C G T      GAA GCC T  T GAT  <-------------------------------------------------
                                     Met
     euGluGlnAsnProLysLeuLysProHisAlaThrThrValPheValMet

                                                                            T             '/          700
Para CTAATGTTAATAAAATAGTATTTTGCCTATTTAAACAACCAAAAATTTAGACACAACTATATAAAACATTTAAATTCTTGTGGTTTATGATACCTTGATC
Kid  ----------------------------------------------------------------------------------------------------

       T                                                      AGTCGCAT                                800
Para TACAATGATTCCAACTTCCCGTGTTGCATTTATGAGTTGTGCTAGCAACAGTCGCAT++++++++CACAGTCGTCTATTCCAGAAAGGACGACTGTGACT
Kid  ----------------------------------------------------------------------------------------------------

                               A                                                           G          900
Para CTTGAGACATATCAAAGCAAAGCTCAGCAATTTTTATGTTTCTCACTTGCTCTGTTCTTTTTCTCTGGTACTTGTCCTGGAAAGACATGTGAGTCTGCGG
Kid  ----------------------------------------------------------------------------------->GT C    T  A  T

                                                                                         ThrCysGluSerAlaV

            C                                                                                        1000
Para TTCAACTTCGGAAAGCCGGAAAAGTGACAGTGAAAGAATCAGACTTGAAAAGAATTGGGGCTATCCACTTCAAAACTGGCGTAGTTAATGAACATTTTGA
Kid  CA        AGC AAT   GC    GTG CTG T CTG  ++++++++++++C    TT         CTC   AA  A   AACG  TCT  G   CT
                                                          100
     alGlnLeuArgLysAlaGlyLysValThrValLysGluSerAspLeuLysArgIleGlyAlaIleHisPheLysThrGlyValValAsnGluHisPheGl

                                                                                                     1100
Para GGTACTACCCTGGCCACTTAGTAGATATAATTCCCTAAGTGTAATCCAAACATTTGTTGTTTAGAGTCAAATTATTATTATTCTGTATGGTGGTTCTTGA
Kid   <--------------------------------------------------------------------------------------------------
     u

                                                                                                     1200
Para ATAATCGATCTTATTATGGTATTTACTAATTATATTATGCATGGGAAAAACGATTTGTAGGTCACAAGGTTTGCACTTTTGGAGACCATAAAGGAAGCAG
Kid  ----------------------------------------------------------->  GGT  AAGAG  TT GC TA A  AT

                                                                 ValThrArgPheAlaLeuLeuGluThrIleLysGluAlaV

                             14mer                     BclI                       17mer          ScaI1300
Para TACCAGAAATGTGGTCACCTGAGATGAAGAACGCATGGGGAGTAGCTTATGATCAGTTGGTTGCTGCCATCAAGTTCGAAATGAAACCCTCCAGTACTTG
Kid   TGG   C AA   A TGA   AC C GC CT  TCTC A C    C  C   G A    CA  A  T  T  AAAG C TATGCTTAGGATT CCT  A
                                                                   150
     alProGluMetTrpSerProGluMetLysAsnAlaTrpGlyValAlaTyrAspGlnLeuValAlaAlaIleLysPheGluMetLysProSerSerThr

                    (C pL305)                                                       pL209>           1400
Para AGAATTTTTATAGTTCTTGGAACAATTGGGTTTGAATAATGTGACAAAACTTATACTTAATTACGTTTGCATGAGAGAGAGGTAATAATTGCATAGTGTA
Kid  T*

                                                                                                     1500
Para TAACTTGCATATGTATCATAGTGTGACGCAATCTCTCCACTTGTGTTGTTCATCTTGTTCAAAAGGAATTAGTCTTTCACTTTACATTTTGGGTGGAAGT

                           pL201>
Para ATGGAATGAAATCAGAGTTTCATTGATT 3'
```

Fig. 4 Comparison of the coding regions of the *P andersonii* haemoglobin gene (sequence indicated is pL209/305, deviations seen in pL201/401 above the sequence) with those of kidney bean haemoglobin[5] Spaces in these regions indicate identity of leghaemoglobin nucleotide sequence to *Parasponia* The kidney bean introns (↔) are different in length, very low in homology and are not compared The same applies to the 5′ and 3′ untranslated regions Also indicated are the positions of oligonucleotides used in cDNA clone identification (14 mer mixture CTC/TTACTTTTTGCG, 17mer mixture TTC/TAAGCTTTACTTTTGG) The derived (and corrected with respect to ref 9) *P andersonii* haemoglobin I amino-acid sequence is given in the bottom line S_1 indicates mapped start of transcription Presumptive promoter elements are underlined (- - - -)

Methods. DNA sequencing was according to Messing[24] following subcloning of appropriate fragments into M13 mp phages Sites used for subcloning of cDNA clones were *Pst*I and *Sau*3A, and for genomic clones were *Eco*RI, *Hpa*I, *Bcl*I and *Sca*I Synthetic oligonucleotides were made to facilitate sequencing of IVS-2 from both strands The DNA sequences of pL201, pL209, pL305 and pL401 were determined at least twice from different M13 mp subclones

has a nitrogen-fixing symbiotic association with *Frankia*, an actinomycete[7] As *Parasponia* is symbiotically associated with *Rhizobium*, this degree of haemoglobin homology suggests separate evolution of plant haemoglobins and nitrogen-fixing symbioses We have also examined DNA from the genus *Trema* (Ulmaceae)[21], a non-nodulating close relative of *Parasponia* The *Parasponia* cDNA probe detects a related sequence in *Trema* at high stringency hybridization conditions, sequence analysis will indicate whether this represents a homologous haemoglobin gene and whether the gene is potentially functional Perhaps haemoglobin genes have survived because of cryptic expression in non-symbiotic tissue in all plants, with abundant expression only in the N_2-fixing nodules of relatively few species scattered through different plant families

We thank Ms G Koci and Ms J Norman for technical assistance, Dr M J Trinick for his supply of nodulated *Parasponia* seedlings, Dr B Barlow for his advice in designing Fig 1 and Drs A and K Marcker for fruitful discussions

Received 22 May, accepted 2 September 1986

1 Tjepkema, J D *Can J Bot* **61**, 2924-2929 (1983)
2 Appleby, C A *A. Rev Pl. Physiol* **35**, 443-478 (1984)
3 Hyldig-Nielsen, J J *et al. Nucleic Acids Res* **10**, 689-701 (1982)
4 Wiborg, O, Hyldig-Nielsen, J J, Jensen, E Ø, Paludan, K & Marcker, K A *Nucleic Acids Res* **10**, 3487-3494 (1982)
5 Lee, J S & Verma, D P S EMBO J **3**, 2745-2752 (1984)
6 Trinick, M J & Galbraith, J *New Phytol.* **86**, 17-26 (1980)
7 Akkermans, A D L & van Dijk, C in *Nitrogen Fixation* Vol 1 (ed Broughton, W J) 57-103 (Clarendon, Oxford, 1981)
8 Appleby, C A, Tjepkema, J D & Trinick, M J *Science* **220**, 951-953 (1983)
9 Kortt, A A, Burns, J E, Trinick, M J & Appleby, C A *FEBS Lett* **180**, 55-60 (1985)
10 Fleming, A I, Wittenberg, J B, Wittenberg, B A, Dudman, W F & Appleby, C A *Biochim. biophys Acta* (in the press)
11 Cronquist, A *An Integrated System of Classification of Flowering Plants* (Columbia University Press, New York, 1981)
12 Dahlgren, R M T *Bot J Linn. Soc.* **80**, 91-124 (1980)
13 Dickerson, R E & Geis, J in *Hemoglobin Structure, Function, Evolution and Pathology*, 68-69 (Benjamin/Cummings, California, 1983)
14 Blanchetot, A, Wilson, V, Wood, D & Jeffreys, A. J *Nature* **301**, 732-734 (1983)
15 Efstratiadis, A *et al. Cell* **21**, 653-668 (1980)
16 Antoine, M & Niessing, J *Nature* **310**, 795-798 (1984)
17 Gō, M *Nature* **291**, 90-92 (1981)
18 Zharkikh, A A, Solovyov, V V & Kolchanov, N A *J molec. Evol.* **21**, 42-53 (1984)
19 Runnegar, B *J molec. Evol.* **21**, 33-41 (1984)
20 Dayhoff, M O, Hunt, L J, McLauchlin, P J & Jones, D D in *Atlas of Protein Sequence and Structure* (ed Dayhoff, M O) 17-30 (National Biomedical Research Foundation, Washington, 1972)
21 Akkermans, A D L, Abdulkadir, S & Trinick, M J *Pl. Soil* **49**, 711-715 (1978)
22 Maniatis, T, Fritsch, E F & Sambrook, J *Molecular Cloning a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982)
23 Gubler, U & Hoffman, B J *Gene* **25**, 263-269 (1983)
24 Messing, J *Meth. Enzymol.* **101**, 20-78 (1983)
25 Hanahan, D *J molec. Biol.* **166**, 557-580 (1983)

New Books
from
CROOM HELM

■ COMPUTER GRAPHICS IN BIOLOGY
Robert Ransom and Raymond Matela, Department of Biology, The Open University
Reviews the uses of computer graphics in teaching, simulation work, and as a tool for analysis and examination of results It considers general methodologies and specific applications such as in neurobiology and molecular modelling
£25 00 hbk 0-7099-4106-4 250 pages September 86

■ DATING AND AGE DETERMINATION OF BIOLOGICAL MATERIALS
Edited by Michael R Zimmerman, Coney Island Hospital, New York, and J Lawrence Angel, Smithsonian Institution, Washington DC
Considers the methods available for determining the age of biological materials in the context of medicine, biology, earth sciences, palaeontology, archaeology and anthropology
£45 00 hbk 0-7099-0470-3 304 pages June 86

■ MODELLING IN BEHAVIOURAL ECOLOGY
Dennis Lendrem, Department of Psychology, University of Newcastle-upon-Tyne
Describes models of animal behaviour and outlines how the modellers came to frame their problem in mathematical form Aimed at the advanced student in ecology and behaviour
£25 00 hbk 0-7099-1691-4 250 pages October 86
£12 95 pbk 0-7099-4119-6

■ ECONOMIC AND APPLIED GEOLOGY
W G Shackleton, South Australian College of Advanced Education
This book provides an excellent general introduction to the subject for undergraduate students of geology, geophysics, mining geology and for civil engineers
£19 95 hbk 0-7099-3387-3 256 pages September 86
£ 8 95 pbk 0-7099-4440-3

■ THE BEHAVIOUR OF TELEOST FISHES
Edited by Tony J Pitcher, School of Animal Biology, University College of North Wales
Reviews major topics in fish behaviour, ranging from genetic and motivational inputs to behaviour, to sensory modalities, behavioural ecology and individual species studies
£42 50 hbk 0-7099-2070-9 560 pages January 86

■ A FUNCTIONAL BIOLOGY OF MARINE GASTROPODS
Roger N Hughes, School of Animal Biology, University College of North Wales, Bangor
'succeeds in portraying much of the excitement of the recent work that has been carried out within the field it can be warmly recommended to all those with more than a passing interest in molluscs and especially those persons running courses on the evolutionary biology of invertebrates' *Times Higher Education Supplement*
£25 00 hbk 0-7099-3746-6 256 pages February 86

■ FEEDBACK AND MOTOR CONTROL IN INVERTEBRATES AND VERTEBRATES
Edited by W J P Barnes, Department of Zoology, and Margaret Gladden, Institute of Physiology, University of Glasgow
Brings together the fields of invertebrate and vertebrate neurobiology and should interest zoologists and physiologists
£40 00 0-7099-32774 480 pages December 85

■ GENE MANIPULATION AND EXPRESSION
Edited by Robert E Glass University of Nottingham and Jaroslav Spizek Czechoslovakian Academy Prague
Presents selected papers from a symposium held in 1984 in Czechoslovakia It covers rearrangement of genetic material and DNA cloning vectors, RNA polymerase, the role of cyclic AMP in gene expression, and translational control
£45 00 hbk 0-7099-3818-7 576 pages January 86

■ GRAZING IN TEMPERATE ECOSYSTEMS: LARGE HERBIVORES AND THE ECOLOGY OF THE NEW FOREST
R J Putman, Department of Biology, University of Southampton
This book combines a general review of our knowledge of the ecology of the New Forest with original research which shows the effects of continued heavy grazing on a temperate ecosystem
£17 95 hbk 0-7099-4036-X 200 pages May 86

■ THE SOCIOBIOLOGY OF ETHNOCENTRISM: EVOLUTIONARY DIMENSIONS OF XENOPHOBIA, DISCRIMINATION, RACISM AND NATIONALISM
Edited by Vernon Reynolds, University of Oxford Vincent Falger University of Utrecht and Ian Vine, University of Bradford, for the European Sociobiological Society
Presents sociobiological studies which focus on the topic of in-group/out-group behaviour
£27 50 0-7099-4222-2 336 pages November 86

■ FISH PHYSIOLOGY: RECENT ADVANCES
Edited by Stefan Nilsson and Suzanne Holmgren, University of Goteborg Sweden
Reviews a number of major advances including physiology of the circulation digestive physiology and physiological toxicology
£27 50 0-7099-1837-2 224 pages July 86

■ A FUNCTIONAL BIOLOGY OF NEMATODES
David A Wharton, Department of Zoology, University of Otago New Zealand
Shows how the basic body plan and life cycle of nematodes has evolved to enable them to exploit a wide range of environments It looks at the phylum as a whole from the viewpoint of intervertebrate physiology, ecology and evolution
£19 95 hbk 0-7099-3602-8 208 pages May 86

■ GULLS AND PLOVERS: THE ECOLOGY AND BEHAVIOUR OF MIXED-SPECIES FEEDING GROUPS
C J Barnard, University of Nottingham and D B A Thompson University of Liverpool
' there is much interesting new detail in the book and some new ways of looking at the interrelationships of flocking and foraging strategies' Nature
£25 00 hbk 0-7099-3230-8 320 pages September 85

■ THE LEAD DEBATE: THE ENVIRONMENT, TOXICOLOGY AND CHILD HEALTH
Edited by Richard Lansdown Chief Psychologist Hospitals for Sick Children London, and William Yule Department of Psychology Institute of Psychiatry London
' the contributors have collected a lot of extremely useful scientific information an important and useful interdisciplinary review Nature
£25 00 hbk 0-7099-1653-1 304 pages April 86
£14 95 pbk 0-7099-1654-X

ORDER THROUGH YOUR BOOKSELLERS
In case of difficulty, send order with payment to the publisher, adding £1 per book for postage and packing

For further information on any title contact Lisa Browne, Publicity Department,
CROOM HELM PUBLISHERS LTD,
Provident House, Burrell Row, Beckenham, Kent BR3 1AT, Tel: 01-658 7813

Reader Service No 38

Oxford
University Press

NEW BOOK

The Particle Explosion

Frank Close and Christine Sutton

This is a spectacular illustrated journey into the subatomic world discovered by physicists during the last hundred years It is the first major illustrated book on particle physics
0 19 851965 6, 240 pp , illus , January 1987 £15

Concepts of Particle Physics: Volume Two

Kurt Gottfried and Victor F. Weisskopf

This authoritative work retraces the material outlined in the first volume, but in far greater detail and with a much higher degree of sophistication
0 19 503393 0, 352 pp , illus , OUP USA, February 1987 £35

Concepts of Particle Physics: Volume One

New in Paperback

Kurt Gottfried and Victor F. Weisskopf

'A splendid review of subnuclear phenomena provides a sophisticated overview without a welter of details
Physics Today on the hardback edition
0 19 504373 1, 180 pp , illus , OUP USA, 1986, paperback £9 95

An Introduction to Liquid Helium

Second edition

J. Wilks and D. S. Betts

The second edition, building strongly on the 1970 edition, brings the reader up to date on liquid ^{4}He, liquid ^{3}He, and mixtures
0 19 851471 9, 180 pp , illus , Clarendon Press, January 1987 £19 95

Theory of Neutron Scattering from Condensed Matter

Volumes 1 and 2

New in Paperback

Stephen W. Lovesey

'Stephen Lovesey has again provided us with a very useful work of reference will be consulted by neutron scatterers for many years'
Nature on the hardback edition
International Series of Monographs on Physics
Volume 1 0 19 852028 X, 352 pp , illus , Clarendon Press, 1986, paperback £17 50
Volume 2 0 19 852029 8, 360 pp , illus , Clarendon Press, 1986, paperback £17 50

The Biological Chemistry of Marine Copepods

Edited by E. D. S. Corner and S. C. M. O'Hara

This is the first book to deal specifically with this topic It is essentially a collection of critical reviews outlining the present state of knowledge and indicating future developments
0 19 854714 5, 360 pp , illus , Clarendon Press, November 1986 £40

The Physics of Semiconductor Devices

Fourth edition

D. A. Fraser

'It is clearly written and is up-to-date I can thoroughly recommend this work' **The Australian Physicist** on the third edition
Oxford Physics Series
0 19 851867 6, 208 pp , illus , Clarendon Press, 1986 £20
0 19 851866 8, paperback £9 95

Principles of Nuclear Magnetic Resonance in One and Two Dimensions

Richard R. Ernst, Geoffrey Bodenhausen, and Alexander Wokaun

A unified view of modern NMR spectroscopy, covering basic principles and techniques for the study of solutions and solids
International Series of Monographs on Chemistry 14
0 19 855629 2, 640 pp , illus , Clarendon Press, November 1986 £60

Patents in Chemistry and Biotechnology

Philip W. Grubb

A new edition of *Patents for Chemists* (OUP 1982) This book guides the lay reader through the complexities of the British, European, and US patent systems
0 19 855222 X, 352 pp , illus , Clarendon Press, November 1986 £25
0 19 855221 1, paperback £12 50

The Thermodynamics and Gas Dynamics of Internal-Combustion Engines

Volume 2

Edited by J. H. Horlock and D. E. Winterbone

A collection of chapters contributed by leading authorities in the field to the memory of R S Benson, author of Volume 1 It is the most comprehensive analytical text available in this subject area
0 19 856212 8, 688 pp , illus , Clarendon Press, November 1986 £75

Nonlinear Ordinary Differential Equations

Second edition

D. W. Jordan and P. Smith

'It is well written, well motivated and contains some recent developments of interest which have not been readily accessible before at this level' **Mathematical Reviews** on the first edition
Oxford Applied Mathematics and Computing Science Series
0 19 859657 X, 400 pp , illus , Clarendon Press, December 1986 £27 50
0 19 859656 1, paperback £14 50

NEW BOOKS

Probability: An Introduction

Geoffrey Grimmett and Dominic Welsh

A concise introduction to probability and random processes at the university level
0 19 853272 5, 224 pp , illus , Clarendon Press, 1986 £20
0 19 853264 4, paperback £9 95

Matrices for Statistics

M. J. R. Healy

This book is intended to cover the necessary ground as briefly as possible Only the simplest of basic mathematics is used Numerical methods for matrices are described
0 19 852207 X, 100 pp , illus , Clarendon Press, November 1986 £12 50

Supercomputers and Their Use

Christopher Lazou

Written in an informal style, this book is aimed at a general, computer literate audience Technical details are kept to a minimum, and are clearly explained whenever they are unavoidable
0 19 853720 4, 240 pp , illus , Clarendon Press, November 1986 £22 50

Subnormal Subgroups of Groups

John C. Lennox and Stewart E. Stonehewer

The object of this book is to give a comprehensive account of subnormal subgroups of both finite and infinite groups, from the early work of Wielandt, to very recent results relating to the elusive submoralizer of a subgroup
Oxford Mathematical Monographs
0 19 853552 X, 320 pp , illus , Clarendon Press, December 1986 £35

Undergraduate Algebra: A First Course

C. W. Norman

The aim of this book is to make the first year student feel thoroughly at home with the theory of rings, fields and vector spaces and especially with the techniques of matrix manipulation
0 19 853249 0, 432 pp , illus , Clarendon Press, 1986 £30
0 19 853248 2, paperback £15

The NAG Library: A Beginner's Guide

Jen Phillips

The aim of this book is two-fold It provides alternative descriptions of a basic selection of NAG routines Secondly, it aims to act as a guide to how a newcomer should approach the *Library*
0 19 853263 6, 256 pp , illus , Clarendon Press, December 1986 £20

Loop Groups

Andrew Pressley and Graeme Segal

A complete and self-contained account of what is known about the subject written from a geometrical and analytical point of view, with quantum field theory very much in mind
Oxford Mathematical Monographs
0 19 853535 X, 328 pp , Clarendon Press, November 1986 £40

Cytogenetics of Mammalian Embryonic Development

A. P. Dyban and V. S. Baranov

A comprehensive review of the role that chromosomes play in normal and abnormal embryonic development
0 19 854584 3, 280 pp , illus , Clarendon Press, December 1986 £35

Blindsight: A Case Study and Implications

L. Weiskrantz

An account of research over a number of years into a particular case of blindsight
Oxford Psychology Series No 12
0 19 852129 4, 200 pp , illus , Clarendon Press, 1986 £19 50

Perception and Communication

New in paperback

D. E. Broadbent

First published in 1958, this book has become recognized as a classic in its field Professor Posner has written a Foreword for this reissue
0 19 852171 5, 352 pp , illus , 1986, paperback £15

The Neuropsychology of Anxiety

New in paperback

J. A. Gray

'The book deserves to be read by all those who are interested in extending neurology into psychiatry'
British Medical Journal on the hardback edition
Oxford Psychology Series
0 19 852127 8, 500 pp , illus , Clarendon Press, December 1986, paperback £15

The Handbook of Ecological Monitoring

Edited by R. Clarke

Written by twelve of the world's experts on ecological monitoring, this book will provide an invaluable introduction to the subject for those requiring a bird's-eye view of the technologies now available
0 19 854590 8, 320 pp , illus , November 1986 £30

Immune Mechanisms in Invertebrate Vectors

Edited by Ann M. Lackie

This volume describes recent ideas and information on methods of recognition and mechanisms of response in the immune systems of insects, crustacea and molluscs
Symposia of the Zoological Society of London No 56
0 19 854004 3, 304 pp , illus , Clarendon Press, November 1986 £40

For further information on Oxford Science Publications, please write to Elizabeth Bone, Oxford University Press, Walton Street, Oxford OX2 6DP

Reader Service No 24

Wiley

CHICHESTER · NEW YORK
BRISBANE · TORONTO · SINGAPORE

NMR OF PROTEINS AND NUCLEIC ACIDS

by **K. Wuthrich,** *Institut fur Molekularbiologie und Biophysik, Zurich, Switzerland*

This comprehensive introduction to both underlying principle and experimental procedures, covers the structure and NMR of biopolymers, including 2D NMR and Nuclear Overhauser effects, resonance assignments and structure determination in proteins, and resonance assignments and structure determination in nucleic acids. Also provided is an up-to-date account of NMR studies for chemists, biochemists, biophysicists and molecular biologists concerned with structure and function of proteins and nucleic acids. The book will enable specialists and non-specialists to evaluate the potentialities and limitations of the method.

0471 82893 9 310pp October'86 £44.60

NATURAL HISTORY OF THE MAJOR HISTOCOMPATIBILITY COMPLEX

by **J. Klein**

This work is a natural history of the major histocompatibility complex written for anyone who is working on the MHC but wants to step back and review the complex in its entirety, for anyone who wants to be introduced to the MHC, or for anyone who is just curious about this much talked-about chromosomal region. By 'natural history' is meant an all encompassing treatment of the MHC, both human and mouse, in its entire range and its details — but presented in a way that anybody can understand without consulting a variety of other sources. Most of the details are summarized in the form of tables whereas the text concentrates on generalizations and discussions.

0471 80953 5 798pp October'86 £90.75

AGE AND SUSCEPTIBILITY TO TOXIC SUBSTANCES

by **E.J. Calabrese,** *Division of Public Health, University of Massachusetts, USA*

Differential susceptibility to pollutant and drug toxicity is a major area of public health concern today. If adverse effects can be prevented from occurring in the high-risk segments, it follows that the general population will also be protected. The causes of differential susceptibility include age, sex, genetic background, nutritional status, the presence of pre-existing diseases, and lifestyle. Previous books by the author have addressed various aspects of these host factors. The present book complements the previous efforts by addressing the age factor in the course of environmental-induced disease.

Series Environmental Science and Technology

0471 84366 0 314pp October'86 £62.25

ECOSYSTEM THEORY AND APPLICATION

edited by **N. Polunin**

The term ecosystem has been widely used, often in a confusing manner since it was first coined some 50 years ago. This book uses the term in the way it was originally defined, as a means of describing the biota which live together as a social unit and also the inorganic and dead parts of the system. Topics covered include arctic desert, marine and tropical ecosystems, their dynamics and structure, economics, evolution and ecological phenomena.

0471 10274 1 462pp August'86 £49.95

SPECTROSCOPY OF BIOLOGICAL SYSTEMS

edited by **R.J.H. Clark,** *University College London* and **R.E. Hester,** *University of York*

This is the first volume of this well-established series, formerly entitled "Advances in Infrared and Raman Spectroscopy", to be devoted to a single theme. The major objectives of the series remain unchanged, ie to integrate theory and practice and to bring together different branches of research in both the academic and industrial spheres through the presentation of critical review articles in fundamental and applied spectroscopy. But its scope has simultaneously been widened beyond the confines of infrared and Raman spectroscopy and focused down more sharply to a well-defined subject area of great current interest. Biological systems represent such an area for applied spectroscopy.

Series Advances in Spectroscopy, Volume 13

0471 90978 5 570pp September'86 £83.00

CALCIUM OF THE CELL

Ciba Foundation Symposium 122

Chairman **P. Baker,** *Department of Physiology, King's College, London*

This international symposium deals with the central role of calcium in intracellular processes that range from muscle contraction to enzyme activation, exocytosis, cell proliferation, and early embryonic development. The symposium should interest research workers in cell biology, cell physiology and molecular biology, it is also relevant to pharmacologists, neurophysiologists and endocrinologists who are concerned with fundamental cell regulatory processes.

Series Ciba Foundation Symposium

0471 91088 0 310pp September'86 £27.50

FOLATES AND PTERINS

Volume 3: Nutritional, Pharmacological and Physiological Aspects

edited by **R.L. Blakley,** *St Jude Children's Research Hospital* and **V.M. Whitehead,** *Montreal Children's Hospital*

This is the first comprehensive review of folates in fifteen years. The contributions to the book provide up-to-date accounts of folate metabolism and chemistry, explaining the metabolic basis for many forms of anemia and providing the basis for understanding the action of certain anticancer drugs, and detail the enzyme targets of some drugs used in treating cancer.

Series Folates and Pterins

0471 81212 9 496pp October'86 £95.45

FUSARIUM SPECIES

Their Biology and Toxicology

by **A.Z. Joffe,** *Hebrew University, Israel*

This clear, in-depth guide describes the Fusarium genus and its toxins — a group of plant pathogens associated with fatal outbreaks of alimentary toxic Aleukia disease in man, and implicated as the active agent in "yellow rain". Noted authority Abraham Joffe presents new and original research into the chemistry, biochemical and toxicological properties, and effects of the four major groups of fusariotoxins. Trichothecenes and Zearalenone and their derivatives, Moniliformin and Butenolide.

0471 82732 0 606pp October'86 £71.75

John Wiley & Sons Limited
Baffins Lane · Chichester · Sussex PO19 1UD · England

Reader Service No 33

Study in indifference

Is the British public not interested in science books? Or is it being poorly served by the review media? John Manger of Oxford University Press on the trans-Atlantic contrasts in the reception accorded to a biography of Albert Einstein.

In 1982 Oxford University Press published *'Subtle is the Lord' ...The Science and the Life of Albert Einstein,* by Abraham Pais. It was judged to be, and remains, a superb piece of scholarship, but not without its problems. Why another Einstein book, asked the doubters, and so soon after the centenary celebration in 1979? We took the plunge anyway, in the belief that we had a truly original and saleable book on our hands. So it turned out to be. Over the next year or so over 30,000 hardback copies were bought world-wide, and subsequently a comparable number of copies of the paperback. The distribution of sales, and how the book was treated by the respective review media in Britain and the United States is illuminating.

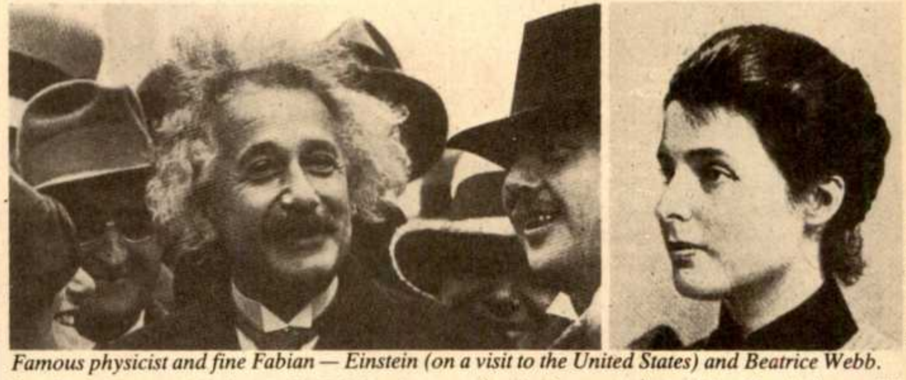
Of the hardbacks, more than 25,000 were sold in the United States, the remainder in the rest of the world, about 2,500 of them in Britain. To put these figures in perspective, an average scholarly physics monograph would be expected to sell 1,000–1,500 copies in hardback world-wide, and more often than not would not be published in paperback. Most new novels sell far fewer than 30,000 copies in hardback.

On both sides of the Atlantic, science magazines and those publications largely or completely devoted to reviews did well by the book. In the United States it was reviewed by *The New York Review of Books, Science, Physics Today, Scientific American* and numerous other weeklies and journals; in Britain by *The Times Literary Supplement, The Times Higher Educational Supplement, Nature* and *New Scientist.*

In the United States, however, the book also attracted much wider attention — from *The New York Times, Los Angeles Times* and *Christian Science Monitor*, as well as from some 30 newspapers across the nation, ranging from the Youngstown (Ohio) *Vindicator* to the Phoenix (Arizona) *Republic*. The review in *The New York Times* was the lead, with picture, in the Sunday books section, perhaps the prime review position in the world. Booksellers, too, were enthusiastic in setting up window displays, while a science book club initially took several thousand copies and then came back for more.

None of the above publicity was particularly exceptional for an American trade publishing house. What was exceptional was the willingness of review editors in such a wide variety of newspapers and magazines to devote space to such a book. It was, after all, not a simple layperson's guide to Einstein, but a scholarly, yet readable, account of the man and his not very simple science.

The book's reception by the "general" media in Britain was in stark contrast. Our approaches to the national weeklies and dailies for review coverage were, in general, met with bewilderment: "Science? We don't have anyone who can review science books"; or "Our readers aren't interested in science"; or "We have 20,000 books a year submitted to us for review, we can't cover them all, you know". The only national newspaper review appeared in *The Observer*, 17 column inches of it. Otherwise, except for BBC coverage in *Kaleidoscope* and on the World Service, the book was virtually ignored. There were no book club sales (there is no scientific book club in Britain).


Famous physicist and fine Fabian — Einstein (on a visit to the United States) and Beatrice Webb.

Mansell Collection

Two questions arise from this experience. Is the British public not interested in science books? Or is it being poorly served by the review media? Bearing on the first of these is the existence in the United States of the well-known and popular "Physics for Poets" course. Students from most disciplines are encouraged to take at least one science option, and physics "without the sums" is widely taught. Stimulating textbooks are produced in large quantities by college publishers and the end result is a sizeable population of students who have a knowledge of space, time, the nature of the Universe, the atom, crystal structure and, yes, even relativity. More importantly, individuals are not encouraged to think it is somehow smart to be scientifically illiterate.

They will certainly also have an awareness of the importance of a character such as Einstein. One of the more amusing responses to our attempts to attract attention to Pais's book was "Well you'd expect the Americans to be interested in Einstein, but he's not really very big over here". Who is big over here? In the week we would have expected reviews of *'Subtle is the Lord'* the dailies and weeklies were obsessed (in terms of review space) by a volume of the diaries of Beatrice Webb. Many column inches, plus suitable period photographs, were devoted to the earnest, but ultimately trivial, meanderings of this fine Fabian. In the great scheme of things, it is difficult to believe that Einstein is not a more important figure to more people, even in Britain, than Beatrice Webb.

The literary editor would doubtless discount this as a somewhat tendentious conclusion, and retort that these kinds of judgements cannot and should not be made when it comes to deciding which books are reviewed in the national press. A literary editor must do what is best for circulation — quite right. But is it a valid assumption that 99 per cent of the readers of the daily and weekly papers have no interest in science for 99 per cent of the time? Because that is the legitimate conclusion to be drawn from the amount of review space given to science books. It should not be forgotten that not only have the social implications of science never been greater, but that there are hundreds of thousands of practising scientists in Britain some of whom can even read newspapers.

Per head of population, the biography of Einstein sold twice as well in the United States as it did in Britain. But it is nonetheless almost certainly true that there is a greater interest in science in Britain than the media lead themselves to believe. The worrying corollary of their indifference (or plain ignorance) is that publishers will tend to take fewer risks, look for fewer new authors and develop fewer new talents if they feel that science books will be "published without trace". The example of Pais's book demonstrates that science books do not have to be trivial to be saleable in global terms. The pity of it is that the review media in Britain remain stubbornly resistant to educating themselves or their readers. □

John Manger is Editorial Director responsible for science and medicine at Clarendon Press, an imprint of Oxford University Press, Walton Street, Oxford OX2 6DP, UK.

One-eyed darwinism

Philip Kitcher

The Blind Watchmaker. By Richard Dawkins. *Longman: 1986. Pp. 332. £12.95. To be published in the United States on 24 November by W.W. Norton, $18.95.*

To Richard Dawkins it is always *the* problem. How are we to account for the complex design in nature, the intricacy of fit between organism and environment that once moved William Paley to propose the analogy of finding a watch on the beach, and thereupon build an argument for the existence of a Creator? Dawkins admires Paley's appreciation of a great mystery, even though he has no sympathy for Paley's theological solution. Thanks to Darwin, we have no need of that hypothesis. The preordained unfolding of the Creator's plan is replaced by the action of natural selection, the blind watchmaker of Dawkins's title.

Dawkins believes, with considerable justice, that the theory of evolution by natural selection is widely misunderstood, and part of his project is to explain that theory in a manner that will demonstrate its power to solve Paley's puzzle. *The Blind Watchmaker* contains passages of exposition that deserve sustained applause. The third chapter begins by distinguishing clearly between single-step selection and cumulative selection, and proceeds through some ingenious and illuminating computer simulations which make plain how cumulative selection can give rise to striking and unanticipated modifications of an original form. It is tempting to suggest that Dawkins's program should be made widely available, to schoolchildren and to eminent astrophysicists, so that the rumbling complaint that compares the effects of selection to a tornado sweeping through a junkyard may finally be laid to rest.

No less brilliant is the discussion at the end of Chapter 4, in which Dawkins the sensitive naturalist replaces Dawkins the clever computer hacker. In 15 pages, he reviews the standard examples of convergent evolution (and some lesser known cases) in prose that makes them extraordinarily vivid. Chapter 5 provides a clear and entertaining account of the elements of contemporary genetics, and Chapter 6 an outline of Cairns-Smith's hypothesis about the origins of life and a useful response to those who claim that the occurrence of life on our planet is so improbable that it signals a miracle. The next two chapters reveal the ways in which natural selection can work constructively to yield "a building up of complexity that has more in common with addition than with subtraction" (p.169). Particularly admirable is Dawkins's lucid presentation of Lande's complex ideas about runaway sexual selection, a discussion in which Dawkins was assisted by Alan Grafen.

The six central chapters of *The Blind Watchmaker* thus constitute a compelling exposition of the main themes of contemporary neo-darwinism, indeed the most compelling presentation of those themes that I know. The stylistic talents that won a broad readership for *The Selfish Gene* are fully in evidence here. But Dawkins aims not only to present what is uncontroversial amongst evolutionary biologists but also to take a stand on matters that continue to spark professional debate. In the end, *The Blind Watchmaker* tries to articulate a particular version of neo-darwinism, and in this part of his enterprise Dawkins is far less successful.

The last three chapters take up three main areas of disagreement in contemporary evolutionary theory. Dawkins considers the proposal that the theory of punctuated equilibria provides an expansion of neo-darwinism, the continuing controversies in systematics, and an assortment of positions that emphasize alternatives to selection as agents of evolutionary change. In this part of the book the real issues are never joined. Dawkins sees clearly that some of the most strident criticisms of neo-darwinism caricature the position that they advertise themselves as attacking. But his own defence is to identify untenable extremes and to portray the version of neo-darwinism that he favours as occupying the sensible middle ground.

I will concentrate on Dawkins's attempt to deflate the punctuationist challenge to neo-darwinism. His strategy here is to analyse the three main claims advanced by the punctuationists, showing that, under analysis, Gould, Eldredge, Stanley and others do not stray very far from neo-darwinian orthodoxy. The success of this strategy depends on how far the analysis proceeds. Eldredge and Gould made a proposal about the geometry of phylogenetic trees: evolution consists of long periods of stasis punctuated by short bursts of speciation. Dawkins fastens on the idea that speciation occurs "suddenly", and argues that the rates envisaged are consonant with those admitted by gradualists. But this is to bypass the main point of difference. Punctuationists contend that evolutionary changes are concentrated around episodes of speciation, and they are (to say the least) sceptical about the importance of either morphological change or speciation within unbroken lineages.

In discussing punctuationist claims about the mechanisms underlying stasis and speciation, Dawkins also finds only what is orthodox. The idea that small molecular substitutions may have large phenotypic effects is, he claims, a part of mainstream neo-darwinism. Thus, the punctuationist emphasis on the importance of modified developmental pathways in speciation can be assimilated into the reigning orthodoxy. The issue here is closely related to that in the chapter where Dawkins considers the complaint that neo-darwinians overemphasize the importance of natural selection in evolution and fail to appreciate the role of developmental constraints.

There, Dawkins concludes by claiming that "all serious Darwinians" would agree that "we can't afford to ignore the constraints on evolution that embryology imposes" (p.311). But although, like Dawkins, most neo-darwinians are now prepared to declare that the power of selection to shape a phenotypic trait is limited by developmental factors, the critics of orthodox neo-darwinism argue that in practice this declaration is forgotten and that explanations of phenotypic traits effectively suppose that natural selection is omnipotent. The debate should centre not on the extreme "Mutationist" who figures in Dawkins's pages, but on the relative importance of various agents of evolutionary change.

Dawkins's failure to identify the real question in this case is reflected in a failure to deal adequately with the punctuationists' account of stasis, and he is led to canvass extraordinary possibilities in interpreting their views. Allegedly, punctuationists believe that species "actively resist change" (p.247), and Dawkins has little trouble in rebutting one obvious reading of this phrase. Charity easily finds alternative interpretations. If there are indeed strong developmental constraints, then throughout most of the history of a species there may be very few options to exercise selection. Only when the environment changes radically, disrupting the mechanisms that typically buffer ontogenesis, will hitherto cryptic variation become apparent and will there be a possibility for evolutionary experimentation. Whether or not scenarios of this kind prevail in the history of life, the claim that they do represents an important challenge to the vision of stately unfolding that permeates orthodox neo-darwinism.

Dawkins's discussion of the third plank in the punctuationists' platform, the thesis that many macroevolutionary trends are the result of processes of species selection, is extremely brief, and the brevity is significant. Instead of presenting his reasons for thinking that species selection is not very important in evolution, Dawkins refers the reader to his previous book, *The Extended Phenotype*. This was a wise decision that would have been advisable at many other places in the last three chapters of *The Blind Watchmaker*, for *The Extended Phenotype* treats the leading controversies in a far more satisfying way than does the new book. However Dawkins does add to his reference another argument that is in tune with the main theme of *The Blind Watchmaker*. Species selection, he contends, cannot account for complex adaptations.

At this point, the sources of the successes and the shortcomings of *The Blind Watchmaker* click into place. Punctuationists will want to combat the argument, alluded to but not presented, for the thesis that species selection is unimportant in evolution, but they may be quite unworried by the charge that species selection cannot account for complex adaptations. Part of their case is that the prevalence of complex adaptations in the history of life has been exaggerated by orthodox neo-darwinians. *The Blind Watchmaker*, which makes the explanation of complex adaptations central to the neo-darwinian enterprise, sets up from the beginning a framework in which the character of the interesting debates within contemporary evolutionary theory cannot be made clear.

The book starts with an attempt to present Paley's puzzle. Dawkins would like his readers to feel the grip of the problem of explaining organic complexity, and he would like to present that problem in a way that does not prefigure the darwinian solution. But how exactly is the notion of a well-designed, complex object to be characterized? Dawkins displays great ingenuity in his effort to avoid the trivializing answer that explicitly defines a complex adaptation as the product of a history of selection. For those who view neo-darwinism as flawed by its overemphasis on adaptation, his struggles are likely to seem hopeless. They will suggest that the problem is to define a class which will turn out to include just those traits whose presence is to be understood in terms of a history of individual selection without making use of ideas that would have been unavailable to Paley (or any other pre-darwinian impressed with the apparent design of organisms). But, claim the anti-adaptationists, cases of apparent design are a motley. Some, but by no means all of them, are to be understood in terms of the action of individual selection. Only by introducing the concept of selection explicitly into our characterization of the class of "well-designed complex objects" can we manage to present Paley's problem in a way that allows for Dawkins's preferred darwinian solution. Hence, Dawkins's emphasis on the evolutionary explanation of "complexity" disguises an unargued commitment to the view that the problem of adaptation is central to evolutionary theory, and a corresponding slighting of important alternatives. For Richard Dawkins — but not for all sophisticated evolutionists and, arguably, not for Darwin — it is always *the* problem.

For all my reservations, I am confident that *The Blind Watchmaker* will do much good. At its heart is a remarkable account of some of the most important ideas in the history of science. Dawkins should clear up many more misapprehensions than he fosters, and his readers should gain a sense of the excitement of contemporary evolutionary biology. If they are moved to read further, then their initial acquisition of prejudices is likely to be corrected. The details, we may hope, will come later. ☐

Philip Kitcher is a Professor in the Department of Philosophy, University of California, San Diego, La Jolla, California 92093, USA, and author of Vaulting Ambition: Sociobiology and the Quest for Human Nature *(MIT Press, 1985).*

NATURAL
HISTORY
MUSEUM
PUBLICATIONS

Mammal Evolution: an illustrated guide
R.J.G. Savage & M.R. Long
Unique colour restorations and other illustrative material combine with Professor Savage's authoritative text to provide a wide-ranging study of these diverse animals.
256pp, 120 colour, 120 b/w illustrations, 300 line drawings.
Hbk. 0 565 00942 7 **£17.50**

The Ancestry of The Vertebrates
R.P.S. Jefferies
The problem of the origin of the vertebrates is solved by deriving them from calcichordates, fossils found in marine rocks of Cambrian to Devonian age.
300pp, 300 b/w illustrations.
Hbk. 0 565 00887 0 **£50.00**

A Manual of Forensic Entomology
K.G.V. Smith
An important new book in this neglected area of entomology. Will be of use to forensic scientists, ecologists and archaeologists.
256pp, 402 figs.
Hbk. 0 565 00990 7 **£17.50**

Jobling's Anatomical Drawings of Biting Flies
Ed. D.J. Lewis
356 of Boris Jobling's drawings are published in this volume for the first time, the originals being held by the Wellcome Museum of Medical Science. Will be of particular interest to medical entomologists.
96pp, 96 plates.
Hardback. 0 565 01004 2 **£9.95**

Sales Department, Natural History Museum Publications, Cromwell Road, London SW7 5BD

Reader Service No.46

What the scientist has to say

Arthur Kornberg

Storm Over Biology: Essays on Science, Sentiment, and Public Policy. By Bernard D. Davis. *Prometheus, Buffalo, New York: 1986. Pp.324. $22.95.*

SCIENTISTS make poor politicians. They have been self-selected for their interest in things rather than people, and they are generally neither literary nor articulate. They form societies to communicate with one another about their arcane findings but, with some exceptions, they are distinguished from most other professional groups by their social isolation and political innocence.

Were the responsibility of scientists limited to progress in such areas as chess or hieroglyphics, their ineptitude in worldly matters might be tolerated with whimsical affection. But the affairs of government and business, our everyday lives and future depend crucially on science and its technological applications. Must we then rely solely on salesmen, actors and lawyers to make the difficult decisions that require informed judgement about scientific questions? What are we to do?

To compound the problem, the media seem to be entranced by the pronouncements of those who lack scientific knowledge and of a few scientists who have ideological biases and crave publicity. The brushfire controversies over drugs, toxic wastes and genetic engineering have coalesced into wider conflagrations of hostility to all technological, scientific and intellectual effort. Legitimate concern about biological warfare has turned into all-out warfare against biology.

Fortunately there are a few informed and responsible scientists who speak out and write clearly on issues involving biology and society. One such is Bernard D. Davis. In *Storm Over Biology* he has collected together 44 essays and shorter pieces on "science, sentiment and public policy". All but one of these have appeared in print in the past ten years, many in obscure places. Trained as a physician, now a scientist with a distinguished record of administration and of research in biochemistry and microbiology, Davis has intimate knowledge of the difficult issues that he has chosen to confront.

Among the matters that get attention are sociobiology, affirmative action in medical schools and genetic engineering. The common theme underlying all of these topics is the need to build on reality and protect science from politicization by the left or the right. In particular, on the perennial debate over biological versus social determinism ("nature versus nurture"), Davis makes cogent arguments for rigorous studies of human genetics and behaviour against wailings by political ideologues about the potential abuses of new knowledge.

In his discussions of the applications of recombinant DNA technology, Davis speaks from broad experience in microbial physiology, genetics and evolution. Recognizing the legitimacy of concern about the hazards of genetic engineering, he offers convincing arguments for the improbability of the scenarios that have most alarmed the public. He further argues for caution in genetic intervention in man, and for recognizing the limits to the possibilities for such intervention. Against specious attacks by Jeremy Rifkin, recently abetted by the editorial board of *The New York Times*, he takes a bold stand for both truth and reason.

A particularly moving, previously unpublished piece deals with the history of the response to his pleas for balance between affirmative action and maintaining academic standards in medical schools. Davis — a child of immigrants, a witness to anti-Semitism, a perennial fighter for liberal causes and the first department chairman in the history of Harvard University to appoint a black to tenure — cannot be accused of insensitivity to the need for social justice. Yet his editorial in the *New England Journal of Medicine*, arguing for reasonable stretching in increasing the opportunities for minority candidates, was presented by the news media as a racist attack, and he was vilified by many colleagues at Harvard Medical School in the hysterical atmosphere that swept through academic faculties during the late 1960s and early 1970s.

I sometimes wished for the dashes of humour that lace Davis's lectures, but the issues in these essays are really serious business. In his foreword, Edward Shils suggests that scientists generally underestimate the depth of the current anti-science movement, and he makes it clear why this important set of statements by Davis will reward the attention of the general public as well as scientists. ☐

Arthur Kornberg is a Professor in the Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA.

Sigmund meets Charles

John R. Durant

The Problem of Altruism: Freudian–Darwinian Solutions. By C. R. Badcock. *Basil Blackwell: 1986. Pp. 206. £14.95, $24.95.*

BOTH natural and social scientists regard Freud's psychoanalytic theories as a somewhat baroque construction of over-elaborate and largely untestable speculations. *The Problem of Altruism* is dedicated to the task of restoring them to their rightful place in the scientific sun. Christopher Badcock is a social scientist who regards conventional social theories of all sorts as hopelessly inadequate to their task. These theories, he argues, being committed to various forms of "holism" and "cultural determinism", are all involved in the futile attempt to explain one social phenomenon in terms of another. What is needed instead, he suggests, is a reconstruction of sociological theory on more reductionist and individualist foundations, and he attempts to lay these foundations by combining psychoanalysis with darwinian social theory, or human sociobiology.

At first sight, this seems a most unpromising strategy. Not only does human sociobiology have a reputation of its own for over-elaboration and untestability, but also it appears to be quite distinct in its approach to the explanation of behaviour. Where freudian theory deals with psychological entities such as the conscious, the pre-conscious and the unconscious minds, and with psychological processes such as projection, regression and repression, darwinian social theory deals with biological entities such as genotypes and phenotypes, and with biological processes such as kin selection and reciprocal altruism. Where is the room for marriage between these two?

Badcock's central argument is that the place where psychoanalysis and sociobiology meet is in the field of "dynamic psychology". For example, the American sociobiologist Robert Trivers has analysed the relationship between parents and their offspring in terms of psychological conflict; and he has debated the possible adaptive advantages of the unconscious mind. In each case, a recent sociobiological analysis appears to mirror an older, psychoanalytic view. Modern darwinists, Badcock suggests, are busy rediscovering and vindicating Freud.

Badcock builds his case for the rehabilitation of Freud around the central theoretical problem of sociobiology, namely altruism. Sociobiologists admit only two major mechanisms for the evolution of behaviour which raises the reproductive fitness of another individual at the expense of that of the actor: reciprocity, and kin selection. Reciprocity, he suggests, may have become an important element in human social behaviour with the development of cooperative big-game hunting; and it may have led to the rapid evolution of complex psychological mechanisms enabling individuals better to practise and detect cheating, including elaborate moral codes and the unconscious. Similarly, kin selection may have promoted the evolution of a basic psychological mechanism of identification; and once in place, this may have permitted the development of anything and everything from masochism to heroic self-sacrifice on the field of battle.

The strength of this freudian–darwinian approach to human behaviour lies in its recognition of the need for a cogent psychological theory to place between the abstract principles of biology, on the one hand, and the concrete realities of human society, on the other. Its weakness, however, lies in the sloppiness of the theoretical constructs that lie at the heart of darwinized psychoanalysis. Take, for example, the central notion of identification. Supposedly, our ancestors evolved a basic ability to identify with close kin in order that they should aid them in circumstances that would promote their own inclusive fitness. Well and good; but for such an ability to evolve, it is necessary that it should be highly discriminating; for much (indeed, most) helping behaviour is detrimental to inclusive fitness and could never evolve. What we need to know, then, is what kind of identification mechanism is being proposed. This we never learn. Instead, we are presented with an allegedly evolved genetic–psychological trait which is so indiscriminate that it leads people to behave altruistically towards virtually anyone and everyone. Under what circumstances could such a dangerous trait as this possibly evolve?

Psychoanalytic theory has always suffered from the problem that its central claims are insufficiently specified to permit the formulation of scientifically testable predictions. Recently, a certain genre of sociobiological literature has rediscovered the superficial appeal of equally vaguely-formulated hypotheses about the roots of human conduct. Having set out to combine the two enterprises, it is perhaps not surprising that Badcock has fallen victim to the trap of explaining both too little and too much. This is a pity, for the ideal of a powerful and a predictive evolutionary psychology is a worthy one. We know that Darwin's basic insights can be applied rigorously to behaviour in general. Perhaps at least some of Freud's ideas could be re-cast in such a way that they, too, could be applied rigorously to human conduct; but this has yet to be shown. In the meantime it has to be said that no genuine "freudian–darwinian solutions" are to hand. ☐

John R. Durant is Staff Tutor in Biological Sciences in the Department for External Studies, University of Oxford, Rewley House, 1 Wellington Square, Oxford OX1 2JA, UK.

Scrotum hokum

Walter Gratzer

The Monkey Gland Affair. By David Hamilton. *Chatto & Windus, London: 1986. Pp.155. £11.95.*

ACCORDING to Dr Johnson, a life protracted is protracted woe. In 1953 the good Pope Pius nevertheless journeyed to Switzerland to receive injections of testicular tissue at the hands of the egregious Dr Niehans. His bid for rejuvenation (though presumably not virility, by which it was generally measured) must have been accounted successful, for shortly afterwards the Mephistopheles in this Faustian contract was elevated to membership of the Pontifical Academy. This occurred two years after Serge Voronoff, in his day the most celebrated and successful exponent of testicle (or, as the euphemism of the time had it, monkey gland) implants, had died, discredited, despised and mocked by the society that had lionized him for two decades. Science does not eradicate human credulity, it merely shifts its focus.

The search for the secret of eternal youth has attracted not only crackpots and mountebanks, but also serious and sometimes even distinguished scientists, whose reason could be selectively suspended. The most illustrious of these was Metchnikoff, who was convinced that aging was caused by bacteria in the gut, which destroyed connective tissue. These, he held, could be driven out by abundant ingestion of yoghourt. Examples nearer to our own time include the view that antioxidants can arrest the processes of decline through the elimination of free radicals, and that lead clothing to exclude cosmic radiation will encompass the same end. None of these theories ever achieved much of a popular following; yet in the 1920s one fancies, there must have been queues of rich, aging voluptuaries in the ante-rooms of "Doc" Brinkley's clinic in Kansas and at the Vienna "Vivarium", their trousers draped over their arms, braces dangling, impatiently awaiting their goat testis implants or unilateral vasectomies — which the learned Professor Steinach was able to assure them would result in the retention of sperm in the testis and thus suffusion of the system with youth and vitality. (This is curiously reminiscent of the principle upheld by the yogis, that the emission of sperm, which they equate with manly virtue, enfeebles the body, so that one must learn to aspirate it back after coition. Many Victorian moralists subscribed to the same opinions, but prescribed abstinence as a corrective.) Steinach treated women by the even less appealing method of X-irradiation of the ovaries.

David Hamilton, who is both a surgeon and a historian of medicine, has written a scholarly, absorbing and vivid account of the history of testicular implantation, which thrived in the first half of this century. For Hamilton, the most interesting figure amongst the *galère* of improbable characters that throng his pages is Voronoff, because he was sincere in his efforts, deluded rather than a charlatan, and within limits rational in his beliefs.

These limits were set by an unfathomed though common ignorance, verging on obscurantism, about the techniques of experimental design and evaluation. The same attitudes of mind are still with us, and they animate homeopathy, psychoanalysis and a range of other peripheral disciplines. The reasoning behind implantation therapy was not new, and amounted to little more than that the testis was the seat of fertility, therefore of youth. However the discovery of endocrine secretions encouraged the plausible hope that foreign testicular tissue might be the source of some benign active princi-

PREACHES FUNDAMENTALISM
How a Famous Surgeon Combines Old-Time Religion and New-Fangled Operations on a Strange Medico-Gospel Farm
PRACTICES GOAT-GLAND SCIENCE
M

From *The Monkey Gland Affair*

The Kansas quack — a feature about "Doc" Brinkley in the New York Evening Journal. *Brinkley is shown holding the first "goat gland baby", born to one of the recipients of a goat testis implant.*

ple. Voronoff became persuaded that a successful xenograft must in general come from a closely related species, not for example from ruminants, and many monkeys were emasculated in the interests of this theory. In the United States prison doctors joined in the game, and were able to go one better: Dr Stanley of San Quentin gaol implanted 30 of his charges with testicles from executed felons and later reported highly satisfactory results from 643 assorted gland transplant operations. All these grafts must of course have been rapidly rejected.

Worse yet was Brinkley, the Kansas quack — unencumbered by any genuine medical qualifications. He chose the goat (an odourless variety), as a species traditionally associated with sexual exuberance, to turn the prematurely enfeebled into satyrs. He and his wife introduced what amounted to a conveyer-belt system of surgery. As a decorative touch, he even injected mercurochrome into the severed end of the vas deferens, to impart a rich colour to the urine and impress the patient. Brinkley owned the hotels that served as convalescent centres, and he set up a radio station to advertise his operation (which was offered in standard and *de luxe* versions), as well as a range of pharmaceutical products. He became extremely rich and popular, and fought a close-run election for State Governor before his luck ran out.

Meanwhile in Paris, Voronoff had married into money on a grand scale and had become a social lion. His second marriage was to a surmised daughter of King Carol of Roumania and the infamous Madame Lupescu. Hamilton makes the point that in the aftermath of the Great War, in which the best had perished and the breeding stock of the upper classes been seriously depleted, and when indeed eugenic theories were widely touted, European society was peculiarly receptive to promises of increased virility and fertility. Voronoff, who was a tireless and prodigiously talented self-publicist, took full advantage of these preoccupations. He also extended his activities to the breeding of sheep, in which he saw great pecuniary promise. The results of Lamarckian (or Lysenkoist) experiments, conducted in Algeria, on the effects of testicular grafts on the performance of rams and the wool-yield of their offspring, were proclaimed a triumph; the British Thoroughbred Breeders Association took fright and announced that grafted racehorses would be struck from the stud-book.

Then Voronoff fell like Lucifer. He was brought down largely through the resolute scepticism of a number of French and English vets, who knew an uncontrolled experiment when they saw one — a circumstance shaming to the medical establishment, which had vacillated feebly throughout Voronoff's years of ascendency. The most tenacious and clear-sighted was Henri Velu, who ran a sheep-breeding station in Morocco, and in the end he outfaced Voronoff and his many powerful supporters. Moreover testosterone had shortly before been isolated by (as Hamilton calls him) a Dr Funk. This was none other than Casimir Funk, the man who gave us the word vitamin and was said to have enraged the Mussolini regime when, learning that Italian prisoners taken in the Abyssinian War were being castrated by their captors, he attempted to negotiate a consignment of the resulting material. Injections of testosterone were found ineffective in the treatment of sexual decline. So began Voronoff's years of eclipse.

David Hamilton has penetrated the shell of Voronoff's formidable public persona, and come to like the man within, self-seeking *poseur* though he quite evidently was. *Tout comprendre*, as they say, *est tout pardonner*. The book is dedicated to Serge-Samuel Voronoff. □

Walter Gratzer is in the Medical Research Council Cell Biophysics Unit, Kings College London (KQC), 26–29 Drury Lane, London WC2B 5RL, UK.

Medical matters

Lewis Thomas

The Oxford Companion to Medicine, Vols I and II. Edited by John Walton, Paul B. Beeson and Ronald Bodley Scott. *Oxford University Press: 1986. Pp.1,524. £55, $95.*

I PICKED up the first of these two volumes at 11 o'clock one evening, and could not lay it down until 11:05. This sounds like the start of an uncordial review, but it means only to say quickly that this new encyclopaedia is not made for steady late-night reading. But for leafing through at random, in spare moments, with nothing particular in mind, *The Oxford Companion to Medicine* contains some unexpected pleasures. All in all, it is an earnest but good-natured piece of work, 1,500 or so pages of what seem rather randomly selected topics, some of them just a few lines of definition, others lengthy essays on one or another aspect of medicine, health, disease, medical practice, and quite a lot on the organization and economics of medicine.

Among the surprises, for me anyway, is a long essay on Doctors as Patients ("doctors have a healthy skepticism about the efficacy of the art they practice"), Doctors as Athletes, Doctors as Missionaries, Doctors in Other Walks of Life (the arts, the law, philosophy, the Church, crime).

Although the editors state in their preface that they hope for a readership made up of intelligent laymen as well as physicians, nurses and other health care professionals, including students, my guess is that the volumes will be most appealing to people who simply like browsing, whatever their background. If, as the editors hope, there is "a plastic surgeon wishing to learn more about the role and responsibilities of a psychiatrist", he will find plenty to read, eleven pages on Psychiatry and another ten on Psychology in Relation to Medicine. The psychiatrist, for his part, will find five pages on Plastic and Maxillofacial Surgery, directly following a short paragraph explaining Plaster of Paris. And any practitioner or layman wanting to know more about Death, Dying and the Hospice Movement, or about Government and Medicine in the UK (and also in the USA), or about Community Medicine, or Dentistry, or Cardiothoracic Surgery, or Medical Museums, or the Medical Services of the Royal Navy, or Women in Medicine, will find pages of essays on these and other broad topics. But, along the way, the casual reader will run into the Opening Snap ("associated with mitral valve opening at the beginning of diastole"), Suffocation ("the interruption of breathing by deprivation of air"), Weakness ("lack of strength, feebleness, ill health"), Neurotoxin ("any agent which is toxic to nervous tissue") and a lot of other items not needing even the one or two lines given them.

Whatever the intended readership, one thing *The Oxford Companion to Medicine* is not: it is not, despite the title, any kind of guide for do-it-yourself-at-home medicine, and the lay person who shells out £55 with this in mind will be disappointed. There is, to my own disappointment, not even much material to enlighten the layman on what to stay away from, in the light of those endless shelves of paperbacks in every airport lounge providing promises by immunity-stimulating diets, magical exercises, trace metals, vitamins, rejuvenating lotions, even books on how to think straight. To be sure, if you go looking you can find some of this, under Fringe Medicine, Cults and Quackery. In the absence of an index (which these two volumes could surely use) you have to be lucky and browsing through the F's to find it, and even so this section deals mostly with the health frauds of the nineteenth century and not enough with the more outrageous follies of this one.

In short, this is a curious, rather unpredictable, almost quirky encyclopaedia. Some will find it useful in spots, interesting for some of the essays. I'm glad to possess it, but I'm afraid that long before my pages became thumbworn my thumbs would be pageworn. □

Lewis Thomas is Professor Emeritus at the Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA, and University Professor at the State University of New York, Stony Brook.

Evolution in the twentieth century

Joe Felsenstein

Sewall Wright and Evolutionary Biology. By William B. Provine. *University of Chicago Press: 1986. Pp. 545. $30, £25.50.*
Evolution: Selected Papers by Sewall Wright. Edited and introduced by William B. Provine. *University of Chicago Press: 1986. Pp. 649. Hbk $70, £59.50; pbk $25, £21.25.*

EVOLUTIONARY theory in the 1960s and 1970s was dominated by optimization, by the search for the quantity evolution would maximize. The mean fitness of a population was not precisely what natural selection would maximize in a Mendelian genetic system, so perhaps if we did enough theory we could find out what was. The picture hanging on our wall was of R.A. Fisher, who had emphasized the importance of natural selection in large populations. The contributions of Sewall Wright were acknowledged by all, but were considered to be of secondary importance.

Theoretical population genetics then saw itself as a branch of applied mathematics, and in need of more mathematical rigour. The only data to which we paid attention were gene frequencies of electrophoretic polymorphisms. The kinds of data which morphologists, palaeontologists and most systematists collected were ignored.

By the end of the 1970s many of these evolutionists working above the population level saw the neo-darwinian synthesis as irrelevant to their needs. Punctuated equilibrium and even neo-lamarckian alternatives gained adherents. The resulting controversies over punctuated patterns of evolution and species selection mechanisms have revived interest in complex fitness surfaces and small populations. The evolutionary views of George Gaylord Simpson and Ernst Mayr, both influenced by Wright, have been important in these debates. All of which has led to a remarkable resurgence of the reputation of Sewall Wright.

It is unusual for the details of a theoretician's work to be so relevant 70 years after his views were formed. It is even more unusual for a major account of his work and its influence to appear in his own lifetime. Provine's thorough and throughly admirable examination of Wright's life and influence, which is accompanied by a very useful collection of Wright's papers on evolution, is the best we have for any recent figure in evolutionary biology.

Sewall Wright was born in 1889, and grew up in Galesburg, Illinois, one of a generation of talented and energetic intellectuals from the prairies of Illinois and Wisconsin. One is reminded of people such as the poet Carl Sandburg and the photographer Edward Steichen; in fact, Sandburg married Steichen's sister, and Wright's father Philip was Sandburg's most influential teacher and first publisher (Sewall set the type and ran the press). Provine gives careful attention to these early years, and to the intellectual atmosphere in which Wright trained; his chapters on Harvard's Bussey Institution and on evolutionary theory after Darwin are worth the price of admission by themselves. Here, I was startled to discover how closely the present controversies over the relative importance of species selection and individual selection were paralleled by controversies at the turn of the century.

Provine also goes into detail — more than seems readable — on the physiological genetics which was Wright's main concern until his retirement. As a student learning two fascinating stories, on mammalian coat colours and mathematical genetics, I did not appreciate how closely they were bound together. One is the story of explaining a seemingly arbitrary array of colours by the concentrations and properties of enzymes. Part of its fascination is evolutionary: the appearance of the same genetic variants in different mammalian species (Himalayan in rabbits, Chinchilla in mice and Siamese in cats, for example). The other extends the genetic system to the population level, making predictions from the necessary regularities of the system of inheritance.

Early days — Sewall Wright in 1922.

Only later did I come to realize that the same person, Sewall Wright, was central to both stories, and that coat-colour genetics was critical to the formation of his views on evolution. A list of Wright's PhD students (pp. 189–190) discloses few who concerned themselves with population genetics, but makes his influence on American mammalian genetics palpable. He wrote major reviews of physiological genetics, and influenced Beadle and Tatum's famous work on *Neurospora* which led to the one gene – one enzyme hypothesis. Although Provine deals chiefly with Wright's evolutionary theories, he recognizes the importance of the work on gene action in coat colour in informing Wright's view of the role of gene interaction in the inheritance of quantitative characters.

After describing Wright's years at Harvard (1912–1915), the US Department of Agriculture (1915–1925) and the University of Chicago (1925–1955), Provine devotes most of the rest of the book to Wright's views on evolution and their influence. There are some surprises. For example, in the course of an extended consideration of Wright's "shifting balance" theory of evolution Provine argues that Wright changed his views substantially in the 1940s. Wright himself denies this, but Provine's evidence, largely drawn from Wright's papers, seems incontrovertible. Before 1940, he points out, systematists believed that differences between species were of little or no adaptive significance. Wright's theory, originally developed with animal breeds in mind, could be applied to natural populations by supposing that random differences arose between species by genetic drift, and that selection between species utilized these to produce a net adaptive response. Provine believes that evolutionary biologists during the 1920s and 1930s did not understand the role of between-group selection in Wright's theory, that they understood him to be saying that genetic drift was the primary agent of evolution.

After 1940, as systematists and population geneticists reinterpreted species differences as being largely adaptive, Wright modified his theory. He brought it closer to the original animal-breeding model by having the random differences arise between local populations rather than between species. Provine finds evidence that this unadvertised change of views was a source of much confusion, playing a role in the famous post-war disputes with Fisher and Ford.

The controversies with Fisher are given careful attention. Provine takes pains to point out the reasons why these two came to disagree so completely about the evolutionary forces operating in nature, despite their agreement on most of the mathematical details. As long as we lack a comprehensive picture of the patterns of gene interaction and population structure in a wide range of species, the views of Fisher and Wright will form opposite poles of microevolutionary theory. One pole emphasizes optimization by natural selection in large populations with additive effects of genes, the other interaction of genes and interaction of natural selection and genetic drift. ►

I was surprised, too, to find how much of Wright's time during the 1930s and 1940s was taken up analysing Theodosius Dobzhansky's data for his famous series of papers on the "Genetics of Natural Populations", and how dependent Dobzhansky was on Wright's suggestions and analyses. Provine's treatment is extensive and illuminating, and indicates the underlying objectives of many of Dobzhansky's papers. Indeed, Provine's book will be an excellent starting point for a study of Dobzhansky's papers; through it the new student in evolutionary biology may be able to appreciate that genetic variation was once much more difficult to find and measure than it is today.

As the title indicates, *Sewall Wright and Evolutionary Biology* concentrates upon Wright's views and influence on evolutionary biology, and is not a general account of Wright's life or science. Although Wright's work on mathematical genetics is mentioned, Provine devotes little space to it. A symptom of this is the omission of the five essential "systems of mating" papers of 1921 from the collection of papers. On the mathematical theory Provine makes a few false steps. After pointing out that there are three logically distinct versions of Wright's concept of a fitness surface, he argues forcefully that these are logically unconnected, that Wright has spread confusion by using two of the notions interchangeably, and that population geneticists mistakenly assume that the two notions are related. Provine is mistaken in this: you really *can* derive the gene frequency adaptive surface from the genotype adaptive surface by taking a weighted average.

Wright's personal life is treated, but only rather briefly and circumspectly. Many population geneticists have Sewall Wright anecdotes, most of which do not appear in Provine's book. I did not find my favourites: "Sewall Wright And The Coffee Cup", "Sewall Wright Penniless In Berkeley" and "Sewall Wright Lectured On Genetic Drift". To his credit Provine does mention the most famous of all, noting both Wright's denial and the numerous people who claim to have seen him absent-mindedly erase a blackboard with one of his guinea pigs.

Twenty years ago I happened to be the person who first suggested to Provine that he should interview Sewall Wright (Provine was surprised to hear that Wright was still alive). Now I'm glad I did, because the result is a remarkable book. As we face the interesting possibility that Wright will be the keynote speaker at his own centennial celebration, his influence on evolutionary biology is greater than it has ever been. We will be listening to that speech with more than just polite attention. □

Joe Felsenstein is a Professor in the Department of Genetics, University of Washington, Seattle, Washington 98195, USA.

Redouté requiem

Sandra Raphael

Wonders of Creation: Natural History Drawings in the British Library. By Ray Desmond. *The British Library: 1986. Pp.240. £25.*

A Redouté Treasury: 468 Watercolours from Les Liliacées of Pierre-Joseph Redouté. Text by Peter and Frances Mallary. *Dent, London/Vendome Press, New York: 1986. Pp.228. £40, $50.*

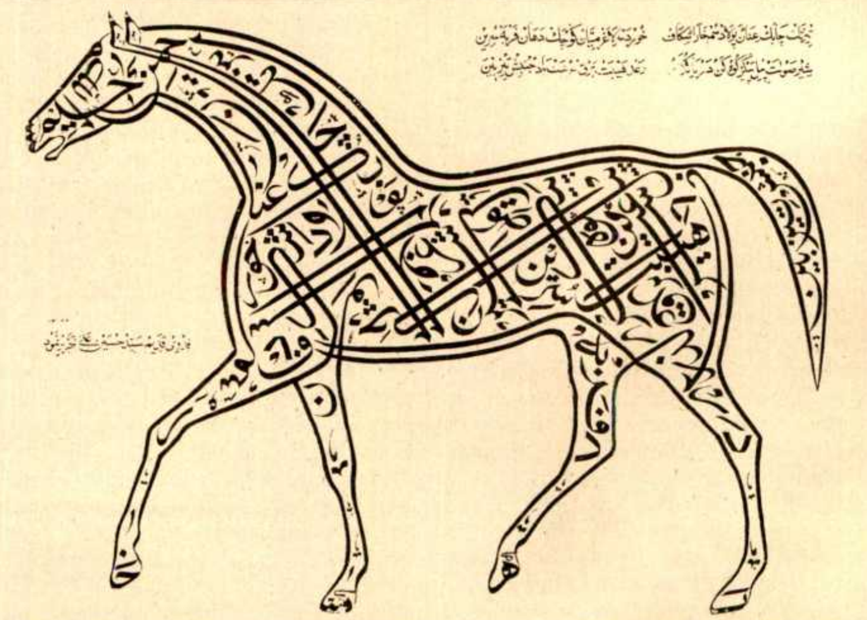
Wonders of Creation contains a selection of illustrations of plants and animals from manuscripts and drawings in various departments of the British Library, with a few pictures from printed books added to demonstrate the changes in style and taste described in the text. The choice seems to have been made almost at random in an attempt to include something for everyone, from flowers and trees, to birds, fishes, snakes, terrapins, rhinoceroses, koala bears, beetles, spiders and Stamford Raffles's "hairy-nosed otter". The sources range from mediaeval manuscripts to nineteenth-century drawings. The coloured plates (nearly 50 of them) are at the back of the book, each with its own short description, while the black-and-white illustrations are used to break up a short text that gives a rather pedestrian survey of the development of illustration as an aid to the descriptive work of naturalists.

The British Library collections in this field are quite rich, but it seems perverse to isolate them from the complementary and even richer hoards in the Natural History Museum in Kensington and the Print Room of the British Museum, still under the same roof as the British Library. In spite of the book's restrictions, the beauty of many of the pictures, oriental or European, will surely please browsers.

On 20 November 1985 Sotheby's in New York sold for $5.5 million the Empress Josephine's copy of Pierre-Joseph Redouté's *Les Liliacées* (1802–1816), 16 volumes containing 468 original watercolours on vellum as well as the printed text. *A Redouté Treasury* is a revised version of the sale catalogue, with colour photographs of the drawings of plants of the lily family and related groups rearranged and supplemented by a longer introductory essay and botanical notes.

Although *Les Liliacées* was sold as a single lot, it has since been dismantled by its buyer, a "breaker" who specializes in taking apart colour-plate books and selling their illustrations piecemeal. This process is sad enough when its subject is a relatively ordinary book, but when it is applied to one that is unique it becomes a bibliographical scandal. Redouté's drawings for *Les Liliacées*, possibly his masterpiece, were bought by his patron Josephine, whose garden at Malmaison provided many models for the artist. The drawings were inherited by Josephine's son, and all but 19 of them were bought by an American dealer, Edward Weyhe, in Zurich in 1935. Now they have been sold again. The drawings have been taken out of their fine bindings, so that this example of a great artist's greatest work has now been scattered and will never again be seen whole. *A Redouté Treasury* may be seen as a requiem for a French national treasure, now ruined. □

Sandra Raphael, 18 Argyle Street, Oxford OX4 1SS, UK, was the natural history editor of the new Supplement to the Oxford English Dictionary, *the last volume of which was published earlier this year.*



Piety in art — an example of Tughra *calligraphy. The opening verses of the* Koran *start at the horse's mouth, while the rest of the text praises the physical qualities of the animal.*

From Wonders of Creation

NEW BOOKS FROM CHAPMAN AND HALL

Tenth Edition

The Biochemistry of the Nucleic Acids

R. L. P. ADAMS, J. T. KNOWLER and D. P. LEADER

The tenth edition of this book has been almost totally rewritten and reorganised to allow a more logical presentation of material — it now provides a clear, wide-ranging and up-to-date account of gene structure and expression. This subject is a core part of all undergraduate courses in biochemistry and molecular biology, and this book is now established as the leading advanced text.

October 1986 544pp
Hb 0 412 27270 9 £35.00 Pb 0 412 27280 6 £16.95

Techniques for the Analysis of Membrane Proteins

Edited by **C. I. RAGAN** and **R. J. CHERRY**

Each chapter of this multi-author volume is derived from first hand experience and describes one approach unique to membrance protein research. Such a broad overview, coupled with specific examples and a guide to the important literature, provides valuable information to researchers wishing to introduce new approaches to their work.

October 1986 456pp Hb 0 412 24970 7 £48.50

The Lipid Handbook

F. D. GUNSTONE, J. L. HARWOOD and F. B. PADLEY

This major reference compendium covers the occurrence, isolation, chemical identification, and technological applications of this large and vital class of natural products.

August 1986 900pp Hb 0 412 24480 2 £95.00

A Topical Dictionary of Statistics

G. L. TIETJEN

A dictionary of statistical terminology in which the contents are arranged by topic rather than alphabetically.

November 1986 200pp Hb 0 412 01201 4 £17.50

Second Edition

Atlas of Stereochemistry – The Supplement

absolute configurations of organic molecules

J. BUCKINGHAM and **R. A. HILL**

This is a supplement to the two-volume Second Edition of the 'Atlas' by Klyne and Buckingham which was published in 1978, and which is the standard reference work on absolute configurations. It extends the literature coverage of the main volumes by six and a half years.

November 1986 c. 336pp Hb 0 412 26000 X £45.00

Third Edition

An Introduction to Radiation Protection

by **ALAN MARTIN** and **SAMUEL A. HARBISON**

Thoroughly revised and up-dated to take account of changes in legislations and standards, this book continues to serve as an introductory text for a wide range of readers involved with radiation in connection with nuclear power, nuclear medicine or for those in research, teaching and industry.

August 1986 264pp
Hb 0 412 27800 6 £15.00 Pb 0 412 27810 3 £7.95

Multivariate Statistical Methods – A Primer

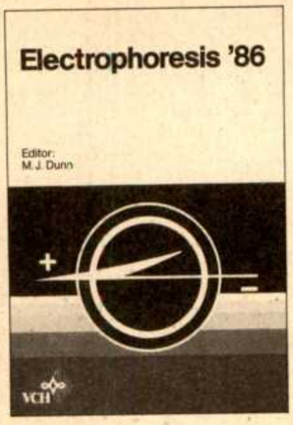
BRYAN F. J. MANLY

This book is an introductory account of multivariate methods for non-mathematicians. The simplest possible approach has been adopted consistent with the aim of conveying the essential principles so that readers can begin to use the methods by themselves.

November 1986 170pp
Hb 0 412 28610 6 £15.00 Pb 0 412 28620 3 £6.95

CHAPMAN AND HALL
11 New Fetter Lane,
London EC4P 4EE.

Reader Service No.31



Electrophoresis '86

Proceedings of the Fifth Meeting of the International Electrophoresis Society, London 1986
edited by M. J. Dunn

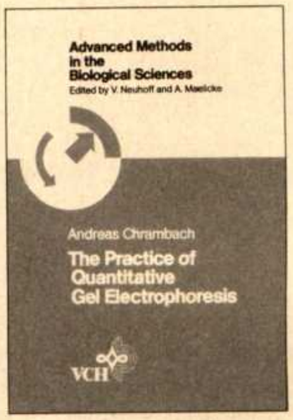
1986. XVIII, 765 pages with 231 figures and 37 tables. Softcover. DM 180.00/US $ 104.20. ISBN 3-527-26566-X

This collection of 72 papers and 72 poster abstracts reflects the latest trends and developments in electrophoretic techniques and their applications. The diversity of the seven major topics covered shows the broad scope as well as the flexibility of these techniques: cell and free flow electrophoresis, isotachophoresis, electrophoresis of nucleic acids, analytical electrophoresis of proteins and other compounds, isoelectric focusing, 2-D PAGE and preparative electrophoresis. Articles on the analytical electrophoresis of proteins and on two-dimensional electrophoresis account for more than half of the contributions, which indicates the focus of today's applications and research.

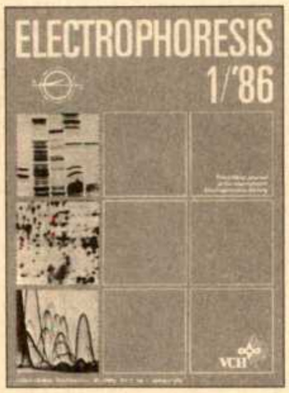
The Practice of Quantitative Gel Electrophoresis

by A. Chrambach

Series: Advanced Methods in the Biological Sciences, Volume 1
edited by V. Neuhoff and A. Maelicke. ISSN 0178-9244
1985. XV, 265 pages with 72 figures and 5 tables. Hardcover.
DM 110.00/US $ 43.60. ISBN 3-527-26039-0



Editorial advisory board of the series:
G. Dowdle, M. J. Dunn, R. Endean, G. Engel, P. Fromageot, M. Osborn, M. Rajewsky, E. Reich, D. Söll, S. Strickland, A. Ullrich, U. Zimmermann.
This book presents gel electrophoretic procedures which are generally applicable to the identification and isolation of charged molecules. All methods have proved reliable and are used in daily laboratory work. Also the equipment required for gel electrophoresis is described in detail. For items that are not commercially available, the reader gets all the information needed to do the construction himself. Theoretical aspects are covered to the extent that they are indispensible for the application of a method. A comprehensive index completes the book and makes it a reliable reference work.



Electrophoresis

The Official Journal of the International Electrophoresis Society

1987. Volume 8. Published monthly. Annual subscription rate including index DM 410.00 plus postage and handling charges. Special subscription rates for personal members of the International Electrophoresis Society, British Electrophoresis Society, French Electrophoresis Society, German Electrophoresis Society, and the Japanese Society of Electrophoresis are available on request. ISSN 0173-0835

Electrophoresis is a journal that stimulates communication between scientists in many fields such as biochemistry, clinical chemistry, genetics, immunology, microbiology and food science by presenting innovative applications of electrophoretic methods.

Electrophoresis publishes articles in English. The highest editorial standards are obtained through the cooperation of an Editorial Board, composed of world-renowned specialists.

Electrophoresis publishes original papers, short communications and review articles. Research reports are written with particular emphasis on advanced methods of electrophoresis. Topics include new and improved analytical and preparative methods and theoretical aspects. Other features include book reviews, important conference dates and miscellaneous information.

VCH

To obtain these products please contact your local bookseller or:
VCH Verlagsgesellschaft, P.O. Box 1260/1280, D-6940 Weinheim, Federal Republic of Germany. Telex 465516 d
VCH Verlags-AG, P.O. Box 151, CH-4106 Therwil. Telex 63475 ch (for Switzerland only)
VCH Publishers, Suite 909, 220 East 23rd Street, New York, NY 10010 (for the USA and Canada only)

Reader Service No.28

The scorbutic story

John Rivers

The History of Scurvy and Vitamin C. By Kenneth J. Carpenter. *Cambridge University Press: 1986. Pp.288. £27.50, $39.50.*

John Le Carré might have told this story better, but I doubt that he could have produced a better book. Like Kenneth Carpenter, Le Carré could have told the tortuous tale of 400 years of medical misunderstanding of scurvy, leading along false trails, down blind alleys and round and round in circles. And Le Carré would if anything have made the reader more aware that, because scurvy was a disease that particularly afflicted armies and especially navies, we are dealing not only with medicine but with grand strategy. The real difference in treatment would come in the last chapter where, the history of scurvy having become the story of the discovery of vitamin C, the author presents a scientific retrospect in which centuries of confusion are put beneath the spotlight of modern knowledge. If Le Carré had unleashed Smiley on these problems, the dénouement would have been far more slick.

Professor Carpenter, however, has been too honest and thorough in his analysis to pretend that all the questions are now resolved. Some, of course, are: the mystery of why the Esquimaux didn't get scurvy, despite eating no vegetables and fruit, is explained by the ascorbic acid content of fresh raw meat. But other problems remain perplexing even today. Why, for example, were so many cases of shipboard scurvy apparently cured by the once-favoured mixture of vinegar and potassium nitrate? The mixture contains no vitamin C; so was it lies, self deception, a placebo effect on an unprecedented scale or something we don't yet understand. The problem is left hanging.

Some certainties fall apart beneath Carpenter's critical analysis. James Lind, for example, doesn't seem anywhere near so epoch-making when Carpenter has finished with him. On the bicentennial of the publication of *A Treatise on Scurvy*, Lind was beatified by scientists as the inventor of the clinical trial and the first truly modern nutritionist. But the Lind to whom Professor Carpenter draws our attention ends up more human, someone who despite his studies on lemon juice as an antiscorbutic believed that cold and damp air actually caused scurvy. By the end of his life, it seems, Lind was a deeply pessimistic man, doubting his own observations which he was unable to replicate in studies at Haslar hospital, where even the controls mysteriously recovered, and concluding in the third edition of his *Treatise* that "enlarged experience must ever evince the fallacy of all positive assertions in the healing art".

Captain James Cook, too, is seen in a more perceptive light. Cook's circumnavigation of the world was free from scurvy not simply because he used Lind's prescribed lemon juice. He took almost every antiscorbutic that anyone recommended, with the result that the trip provided no test of Lind's hypothesis. Indeed, says Carpenter, Cook may have delayed the general introduction of lemon juice as a prophylactic because every theorist drew sustenance for their pet idea from his provisioning strategy; the award of the Royal Society's Copley medal to Cook, for his success in avoiding scurvy, seems to have been supported by Sir John Pringle, the President of the Royal Society, because he regarded Cook's triumph as evidence in favour of his favourite hypothesis that malted grain was the preventative of choice.

Such re-evaluations are the result of careful research on Professor Carpenter's part, not idle iconoclasticism. Indeed, in springing to the defence of the Admiralty, Carpenter attacks the iconoclasticism of others. It is often held that the continuance of scurvy in the British navies during the nineteenth century was the result of the Admiralty's penny-pinching decision to abandon lemon juice in favour of lime juice from the West Indian colonies which, though cheaper, had no vitamin C activity. The truth is not so simple. Not only are limes a good source of vitamin C, but the words lemon and lime were used interchangeably and it is not always possible to be sure what was being done. Carpenter suggests that the ineffectiveness of lime juice may have been due to catalysis of the oxidation of the ascorbic acid by copper in the stills and vessels used to concentrate the juice, compounded by overlong storage. The Admiralty was doing its best; it was the scientists who were wrong.

This is an extremely detailed book, and at times I wondered whether Carpenter should not have begun with the modern knowledge of vitamin C to aid the reader in interpreting the sheer mass of erroneous evidence and theory that runs down the centuries: a sort of answers first, questions afterwards strategy. But the lack of certainty in the final clarification shows that we don't yet have all the answers, and some questions must just remain. That alone makes the book worthwhile, as well as delightful to read, and a volume that will inform and entertain both historians and scientists — and, indeed, anyone who likes a good mystery story, even if it isn't written by Le Carré. ☐

John Rivers is a Senior Lecturer in the Department of Human Nutrition, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.

INSTITUTE OF TERRESTRIAL ECOLOGY

Landscape changes in Britain
C.J. Barr, C.B. Benefield, R.G.H. Bunce, H.A. Ridsdale & M. Whittaker
A largely visual presentation of results from 2 landscape surveys of Great Britain.
28pp, 0 904282 95 3 **£3.00**

Pollution in Cumbria
P. Ineson
Proceedings of a symposium concerned with acid rain, radioactivity, industrial and agricultural pollution.
92pp, 0 904282 96 1 **£5.75**

Coast dune management guide
D.S. Ranwell & R. Boar
How to use vegetation for coast protection activities, with particular reference to sand dunes.
105pp, 0 904282 93 7 **£6.00**

Distribution and status of bats in Europe
R.E. Stebbings & F. Griffith
31 species are described under the following headings: distribution, habitat, population, threats, conservation measures, bibliography.
142pp. 0 904282 94 5 **£5.00**

Full list from:
Publications Section
Merlewood Research Station
Grange-over-Sands
Cumbria LA11 6JU, UK

Many earlier titles now reduced in price.

Natural
Environment
Research
Council

Reader Service No.20

Inside development

John C. Gerhart

Molecular Biology of Development. Volume L, Cold Spring Harbor Symposia on Quantitative Biology. *Cold Spring Harbor Laboratory: 1986. Pp. 920. Hbk $140; pbk $70.*

AT THE very least, this is yet another outstanding volume to arise from the annual Cold Spring Harbor Symposium, now in its fiftieth year. The papers — over 100 of them — come from leading molecular biological and genetic laboratories; they cover the subjects of nucleo-cytoplasmic interactions, cell lineage, segmentation, homoeotic mutants, homoeoboxes, tissue-specific gene expression, sex determination, cellular differentiation and developmental neurobiology. There is probably no book on the molecular biology of development to match this one, not only in the comprehensiveness of the coverage, but also in the care of preparation of individual papers, an indication of the authors' high regard for the occasion of a Cold Spring Harbor meeting. Most of the contributions are no-nonsense research accounts which assume a technical background and discuss only the immediate implications of results. Broader orientation is available by way of a general introduction by John Gurdon, and a summary of symposium highlights by Gerald Rubin. In this rapidly changing field, the papers cannot be expected to remain current for very long; the volume should be studied forthwith.

What is the molecular biology of development, as revealed in this volume? Joe Sambrook writes in his foreword:

> Developmental biologists are an audacious breed. Their ambition is to provide an intellectually satisfying account of the forces that guide development of multicellular organisms around the circle from fertilized egg to adult to gamete. These processes, however, lie close to the border beyond which genetics may not be feasible nor necessarily useful; furthermore, meaningful biochemical studies pose extraordinary challenges because of the small amounts of material that are available and the general rapidity of events in the early embryo. In the face of these difficulties it is not surprising that for many years developmental biology was essentially a branch of anatomy, in that it was exclusively descriptive in nature. Accurate and elegant though these descriptions were, they led to an understanding of developmental processes less frequently than to an amazement at their beauty.
>
> In recent years, however, developmental biology has undergone a dramatic change and has matured from a descriptive to an analytical science. There is no doubt in my mind that this change stems almost entirely from two technical advances — the ability to use molecular cloning to isolate and characterize mutant and wild-type versions of genes that control or are expressed at specific developmental stages and second, the ability to generate transgenic organisms in which the expression of the introduced gene(s) is correct both spatially and temporally. In consequence, the developmental biologist has not only the capacity to describe but also now to analyze and influence the events that guide a fertilized egg to its destiny.

In the book, it seemed to me that the excellent description by I. Herskowitz of the mating-type locus in yeast approaches most closely the kind of "intellectually satisfying account" now aspired to by molecular biologists studying development (for example those working on sex determination, segmentation and homoeosis in *Drosophila*). Herskowitz indicates (1) how certain master regulatory genes encode DNA-binding proteins which directly control the expression of batteries of cell-type-specific genes, and (2) how these master regulatory genes are themselves regulated by products of the *HO* gene, which is activated within mother cells by the products of yet other genes of the *SWI* family. Finally he suggests (p. 572),

> returning to the egg, we can imagine that regulatory proteins such as the *SWI* products become asymmetrically distributed to sister blastomeres. They activate expression of genes analogous to *HO* that govern activity of a regulatory locus (analogous to *MAT*) that ultimately triggers cell specialization.

Several contributors trace similar thoughts to those of T. H. Morgan (*Embryology and Genetics*, 1934, p. 10):

> The initial differences in the protoplasmic regions [of the egg] may be supposed to affect the activity of genes. The genes will then in turn affect the protoplasm, which will start a new series of reciprocal interactions. In this way we can picture to ourselves the gradual elaboration and differentiation of the various regions of the embryo.

Thus, in scanning this volume, I sensed fair agreement among participants that new information on genetic regulatory hierarchies is compatible with the long-appreciated phenomenon of cytoplasmic localization (an *intra*cellular, lineage-specific phenomenon) and can be grafted on to it to give an intellectually satisfying account now identifiable as the molecular biology of development.

Is there more? As John Gurdon points out, vertebrate development is thought to depend primarily on arborizing cascades of ever more local inductive cell–cell interactions. The highly regulative nature of vertebrate embryos indicates the ongoing ability of individual cells to adapt their paths of development to *inter*cellular conditions. Imaginal disc regeneration and transdetermination also reveal the importance of continuous cell–cell interactions in invertebrate (*Drosophila*) development. These complex, multiple interactions, while mentioned in the volume, have an uncertain relationship to the genetic control of cell differentiation and to lineage-restricted cytoplasmic factors. If embryonic cells produce and respond to a great variety of *inter*cellular contexts, and if a cell's particular path of development (including its changing profile of gene expression) is highly adaptable to that context, then regulatory genes would themselves be subject to extensive regulation, and would seem to be rather less masterful in their control of development, even though they would function to control other genes and play an important inertial rôle in connecting the cell's past to its present. In organisms making use of cell–cell interactions to select developmental paths, gene-regulating genes might be elements of much larger intercellular regulatory circuits having many components, many inputs, multiple states and adaptability. (It's one thing to say that a gene controls other genes; it's another to say that a gene controls development.)

If this is the prospect, then with the help and stimulus of the exciting new analytical techniques, and with genetic dissection, we can look forward to a long period of amazement at the beauty of our enlarging description of development, as the molecular biology of development turns outwards from the nucleus. ☐

John C. Gerhart is a Professor in the Department of Molecular Biology, University of California, Berkeley, California 94720, USA.

NUCLEAR POWER AND ITS ENVIRONMENTAL EFFECTS

For the Layman, Nuclear Technician, Manager

Places the effects of possible nuclear radiation in perspective. Meets the need for a better understanding of radiation and its consequences especially after the TMI accident.

All vital aspects of nuclear power generation are considered, from fundamental principles to plant licensing, from uranium mining to disposal of waste. The safe operation of nuclear power plants, to prevent the escape of dangerous amounts of materials in case of an accident, is fully covered.

By Samuel Glasstone
Walter H. Jordan
American Nuclear Society
555 No. Kensington Ave.
La Grange Park, IL 60525
USA

408 pages
ORDER TODAY
$18.95 Softbound
$26.95 Hardbound

The source for the latest knowledge in Nuclear

Reader Service No.10

Feeling for nature

Robert M. May

Gilbert White: A Biography of the Author of The Natural History of Selborne. By Richard Mabey. *Century Hutchinson, London:1986. Pp.239. £14.95.*

ONLY three books in the English language have gone through more editions than Gilbert White's *The Natural History of Selborne*. This claim leaps out from the blurb on the back of the 1977 Penguin edition, and has left me wondering what ranks third behind the Bible and Shakespeare; probably *Pilgrim's Progress*. The separate editions of *Selborne*, over 200 of them, have appeared at a steady rate since the book was first published in 1789 (note the impending bicentenary), forming an interesting record of changing tastes in book design and changing visions of the countryside as revealed in the illustrations.

On first acquaintance, *Selborne* is a surprising book to have evoked such unwavering affection. As Mabey observes, it is

> the kind of book presented on prize-giving days, ... a work, in all senses, of the old school. Even when I eventually came to read it, I cannot say my opinion changed dramatically. I could not cope at first with its rambling disorder, its sudden plunges into thickets of taxonomic Latin, and, for a while, I failed to notice the feeling behind the often dispassionate prose.

Selborne is organized as a collection of letters, similar to many other works about local natural history by enthusiastic parsons. It endures while these other books ornament the libraries in stately homes because beneath its unassuming surface is a fully contemporary appreciation — quite new in its own time — of plants and animals as more than curiosities, but as parts of a community, interacting among themselves, with their environment and with humans. White raises general questions (such as how populations are regulated), but always against a background of keenly observed detail (the breeding success of swifts, the singing of crickets, or the way fly-catchers keep their fledglings cool in summer by fanning their wings above the nest).

Richard Mabey has produced the first biography of Gilbert White for over 80 years. The difficulty is that not a lot is known about White, except that he lived an uneventful bachelor life in the parish and rarely left it (as Richard Boston observed, reviewing the book in *The Guardian*, how else could White have described the South Downs as "that chain of majestic mountains"). Mabey's book is more in the nature of an introduction to, and appreciation of, *Selborne*, than a biography in the strict sense. At such, it does a fine job of relating Gilbert White and his book to the village of Selborne as it was then and is now. Despite his observations, White never fully gave up the contemporary notion that swifts and swallows disappeared in winter because they hibernated; Mabey empathizes, suggesting that White simply could not imagine why anyone could want to leave Selborne.

Mabey explains that "it was seeing Selborne itself, the source of White's inspiration, that changed my view [quoted above] decisively, and helped me to understand what the book was truly about". The patterns of field and woodland around the village have changed little since White's time, and the village itself conducts its Gilbert White trade in a low-key way (the stained glass memorial window in the church, with White as Saint Francis among his birds, is alone worth a visit: see the cover illustration of *Nature*, 22 December 1983). Visit Selborne and read Mabey. Then read *The Natural History of Selborne*. ☐

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544, USA.

In the groove

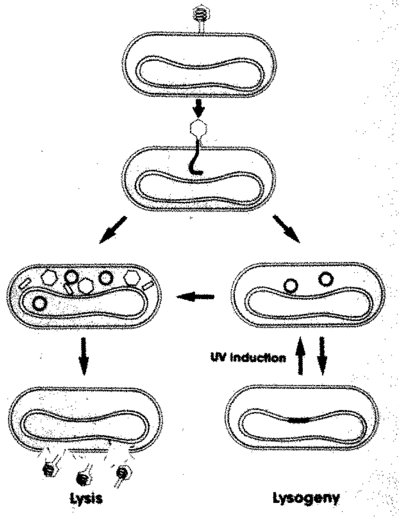
Sydney Brenner

A Genetic Switch: Gene Control and Phage λ. By Mark Ptashne. *Cell Press, Cambridge, Massachusetts/Blackwell Scientific: 1986. Pp.128. Pbk $16.95, £14.50.*

MAX Delbrück did not believe in lysogeny. He thought that the persistence of bacteriophages in some bacterial cultures was due to cryptic infection and he dismissed it as a fringe phenomenon. Everybody he influenced therefore worked on virulent bacteriophages, especially T4, and their work led to one branch of molecular genetics and to the discovery of the fine structure of the gene, the general nature of the genetic code and messenger RNA. Fortunately, Delbrück's influence was not totally pervasive and in Paris, where there was a long tradition of research on lysogenic bacteria, the work of Lwoff and of Jacob and Wollman on bacteriophage λ revealed that lysogeny is in fact due to a gene regulatory mechanism that allows the phage DNA to remain silent in the bacterial chromosome until activated by, for example, ultraviolet light. Their major step was the realization that the properties of the phage mutants on which they worked were isomorphous with those affecting the expression of β-galactosidase, and that in both cases gene regulation was achieved by the negative effect of repression.

Mark Ptashne isolated the λ repressor in 1967 and has spent nearly 20 years in pursuing its molecular structure and mode of action. This little book is an account of his research, focused especially on how the λ repressor and *Cro*, another regulatory protein, form a switch by their interaction with an operator, O_R. In one state, corresponding to stable prophage, the binding of λ repressor inhibits the rightward transcription from the promotor site P_R and stimulates its own production by transcription in the direction from another promotor, P_{RM}. When the repressor is removed by proteolysis induced by ultraviolet light, rightward transcription is turned on and *Cro* is produced. The binding of *Cro* to O_R prevents any further synthesis at P_{RM}, thus turning off repressor, irreversibly throwing the switch and sending the phage down the lytic pathway. Ptashne shows how these events are precisely conditioned by the mode of binding of repressor and *Cro* to the three-component sequence of the operator. The structure of these proteins is known and both recognize the major groove of DNA through a characteristic α-helix–bend–α-helix, a motif now known to be present in a large number of DNA-binding proteins.

This work is a paradigm for the mechanisms of regulation but Ptashne would like to go further and sees the λ switch as the model for switching between develop-



Lysis or lysogeny — the alternative modes of growth of phage λ which are controlled by a switch mechanism operated by λ repressor and Cro *regulatory protein. Once it has entered the host bacterium (top), the phage DNA can either replicate autonomously and then lyse the host; or it can insert itself into the host chromosome and be propagated along with it until the lysogeny decision is reversed by, for example, irradiation of the bacterium with ultraviolet light. (Reproduced, reduced in size, from* A Genetic Switch; *the original is in two-colour.)*

mental pathways in higher organisms. In a general sense this is true because, logically, the λ system appropriately combines positive and negative feedback in a simple way to implement a switch; but this is not necessarily the only way it can be done. Years ago Novick performed a most illuminating experiment. Normally the galactosidase of *Escherichia coli* is repressed, but it can be released from repression (induced) in the presence of lactose or an analogue thereof. Novick briefly exposed *E.coli* cultures to a high concentration of an inducer which was then diluted a thousandfold to provide a maintenance level. In a fraction of the cells, enough of the inducer got in to derepress the genes; the rest remained uninduced. The resulting mixed population maintained itself indefinitely as long as the low level of inducer was present, because induction entails the production of a permease that increases the uptake of the inducer so that even at low levels enough can enter the cells to maintain the induced state. Thus it could be said that *E.coli* had "differentiated" into two cell types with heritable states. This result depends on the properties of the permease: only cells with permease can take up inducer at the low maintenance level and only such cells can have permease induced; transient induction at high levels simply initiates the process in part of the population. Here the logical elements of positive and negative feedback are implemented in another way and it is arguably a better paradigm for some developmental processes than the λ switch.

The analogy between the λ switch and developmental processes should not be carried too far. In discussing the difference between "commitment" and "differentiation" of cells, Ptashne writes

> By analogy, a λ-lysogen is indistinguishable from an uninfected bacterium by casual inspection, but the former and not the latter is 'committed' to lyse and release phage when it encounters an inducing signal in the environment.

However, he quickly realizes that this is not a good analogy because the difference between the bacteria is a *genetic* one and he continues

> The lysogen, of course, has genes that are lacking in the non-lysogen, but it is not difficult to imagine analogous forms of commitment involving only differential gene expression. To take a simple example, the regulatory proteins of a cell might turn on a gene encoding a hormone receptor.

Do we hear a faint echo of the permease paradigm?

But this is quite a minor criticism of a clear exposition of a series of triumphant experiments and insights that have brought us close to the molecular heart of the matter of genetic regulation. ☐

Sydney Brenner is Head of the Medical Research Council Molecular Genetics Unit, Hills Road, Cambridge CB2 2QH, UK.

Eastern originality

Christopher Cullen

China: Land of Discovery and Invention. By Robert K. G. Temple. *Patrick Stephens, Wellingborough, UK: 1986. Pp.248. £12.95.**

READERS of this journal will need no more than a brief reminder of the importance of the work of Joseph Needham. With the help of a large group of collaborators, over the past 40 years he has been engaged in a great project of research and publication on the history of science, technology and medicine in pre-modern China. The results are appearing in the pages of *Science and Civilisation in China*, the seven volumes of which are being published by Cambridge University Press as some 30 separate tomes (about half of them have appeared to date). This work is widely regarded as the most influential British contribution to historical scholarship in the present century, and it has certainly completely changed the perspective of many well-informed Westerners on the origin of important aspects of modern science.

As the weight of his research has accumulated, Needham has had to move far from his original aim of writing a short book for the general reader. Robert Temple's book, which is compiled almost entirely on the basis of material already published or awaiting publication in Needham's volumes, is intended to meet this need. Temple has divided his account into 11 broad subject areas, such as "Agriculture, Medicine and Health", "Magnetism" and "Warfare". Under these he has grouped a hundred short, illustrated essays on particular topics. Temple is clearly no mere hack compiler of coffee-table books: his writing is clear and vigorous, and he succeeds in communicating his obvious enthusiasm for the subject. It is hard to imagine anyone with a general interest in the history of science failing to find a great deal in this book to intrigue and delight.

There are some flaws. Temple makes no systematic attempt at writing history; the book is a series of vignettes on "the wonders of Chinese science", and as such does not convey any sense of the social and philosophical background from which Chinese science, technology and medicine arose and against which they functioned. The choice of topics and the balance between them sometimes seems odd: the section on engineering devotes more space to the novelties known as "spouting bowls" than to cast iron. Some of the statements are out of date or just plain wrong, and the one-page list of "further reading" is so incomplete that it does not even mention Colin Ronan's readable and reasonably priced abridgement of Needham's work, also published by Cambridge University Press. On the whole, however, Temple deserves the gratitude of his readers for having provided a ready means of access to a fascinating subject. ☐

Christopher Cullen, 86 Mount View Road, London N4 4JX, UK, is a former Research Fellow of Clare Hall, Cambridge, and has published mainly in the field of ancient Chinese astronomy and mathematics.

---

*To be published in the United States in January by Simon & Schuster, under the title *The Genius of China: 3000 Years of Science, Discovery, and Invention*. Price will be $19.95.

From *China: Land of Discovery and Invention*

Sowing sophistication — spring wheat being sown by seed drill in northern China. Seed drills, which are much less wasteful than broadcasting seed by hand, were available in China from the second century BC. *It was almost 2,000 years later that they came into use in Europe.*

Major new titles from Gordon & Breach and Harwood

New Perspectives in Nuclear Medicine

Proceedings of the IX Congress of the Latin American Association of Societies of Nuclear Medicine and Biology (ALASBIMN), Montevideo, Uruguay, December 9-13, 1984

Edited by Peter H. Cox and E. Touya
Volume 2 in the Monographs in the Nuclear Medicine series

A valuable reference work, reporting on the discovery and applications of new nuclear methods in diagnosis and therapy that have been made possible by recent technological advances in the field. Part 1 contains clinical *in vivo* studies; part 2 reviews instrumentation, laboratory investigations and *in vitro* studies.

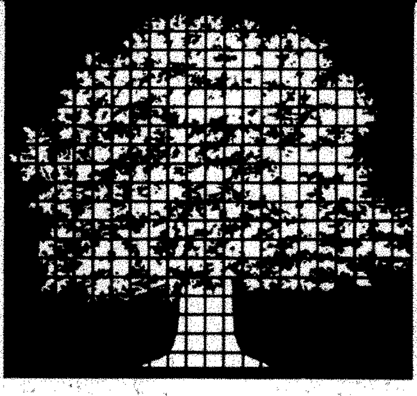
Tentative publication date:
January 1987 approx. 822 pp.
*2 volume set, hardcover 2-88124-131-X: $160.00; SAS member price: $80.00**
A Gordon & Breach set

New Approaches in Cardiac Mechanics

Edited by Kazuo Kitamora, Hiroshi Abe, and Kiichi Sagawa

This volume contains the results of collaborative research by cardiologists, physicists and engineers into cardiac mechanics. Emphasising the research and development of relevant methodologies, the articles review new approaches to cardiac function, from the mode of muscle contraction and its control to clinical assessment.

May 1986 330 + ix pp.
Hardcover 2-88124-126-3: $80.00
A Gordon & Breach book published jointly with Japan Scientific Societies Press.



Higher Integrative Systems of the Brain

by A.S. Batuyev

For physiologists, psychologists and philosophers, this book reviews the latest experimental data and ideas, and summarises current knowledge on the physiology of the brain's higher integrative systems and their role in shaping complex behavioral patterns in man and other animals.

Tentative publication date:
May 1987 approx. 286 pp.
Hardcover 2-88124-143-3, tentative price: $82.00
A Gordon & Breach book

Polymer Yearbook 3

Edited by Richard A. Pethrick

". . . the comprehensive academic library will want to have the series for its annual coverage of polymer science, not otherwise available to an interested reader. The layout, printing, accuracy, coverage and overall quality are excellent." – *Choice*

August 1986 337 + xiv pp.
Hardcover 3-7186-0342-X: $80.00
Softcover 3-7186-0341-1: $30.00
A Harwood book

Essays on DNA in Clinical Medicine

Edited by T.J. Wilkin

Discussing molecular biology techniques and their applications in areas from hormone bioassays to the diagnosis of genetic disease, and the ethical issues that arise from the use of bioengineering, this book offers the health practitioner, hospital clinician and medical student a solid understanding of some of the most exciting developments in modern medicine.

Tentative publication date:
March 1987 approx. 180 pp.
Hardcover 3-7186-0360-8, tentative price: $30.00; Softcover 3-7186-0361-6, tentative price: $18.00
A Harwood book

Synthetic Membranes

Edited by Maynard B. Chenoweth
Volume 5 in the MMI Press Symposium series

Covering the chemical, physical and commercial aspects of synthetic membranes, this book discusses the state of the art in membrane technology, its applications to medicine and areas such as food processing and water purification.

August 1986 288 + vii pp.
Hardcover 3-7186-0327-6: $80.00
*SAS member price $30.00**
A Harwood book

* **The Science and Arts Society (SAS)** is a bookclub designed to allow individuals to purchase high-level reference books with a saving of around 40% on all prices. Most Gordon & Breach and Harwood titles are included as well as books and special offers from other publishers from time to time. To become a member, simply include a $5.00 lifetime membership fee with your payment.

Order these Gordon & Breach and Harwood titles through:
STBS
P.O. Box 786 Cooper Station, New York, NY 10276, USA · P.O. Box 197, London WC2E 9PX, UK

Reader Service No.14

NEW AUTUMN TITLES FROM

·ACADEMIC PRESS·

Advances in Drug Research *Volume 15*

Edited by **B. Testa**
November 1986, 256pp., $54.00/£30.00(UK only),
ISBN: 0.12.013315.6

Contents: Toxication Mechanisms in Drug Metabolism, *H.-G. Neumann.* Elements for the Rational Design of Peptide Drugs, *J.-L. Fauchere.* Tissue and Receptor Selectivity: Similarity and Differences, *T.P. Kenakin.* Inhibition of Prostaglandin, Thromboxane and Leukotriene Biosynthesis, *J.A. Salmon.* Foodstuffs as Sources of Psychoactive Amines and Their Precursors: Content, Signifiance and Identification, *R.T. Coutts, G.B. Baker and F.M. Pascutto.*

Special Publications of the Society for General Microbiology No. 18

Bioactive Microbial Products 3

Downstream Processing

Edited by **P.J. Bailey, J.D. Stowell, and D.J. Winstanley**
November 1986, 340pp.,
$35.50/£19.50(UK only), ISBN: 0.12.672961.1 Pbk.
$63.00/£35.00(UK only), ISBN: 0.12.672960.3 Hbk.

Contents: Introductory Remarks, *M.D. Lilly.* Significance of Downstream Processing, *B. Atkinson.* Industrial Approaches to Fermentation Recovery R&D, *K.G. Taksen.* High Level Microbial Expression and Purification of Recombinant Proteins, *T. Atkinson, D.A. Barstow, J. Court, N.P. Minton, R.J. Sharp and R. Sherwood.* A Review of Antibiotic Isolation Techniques, *R. Carrington.* Recovery of Antibiotics Using Column Extraction Methods, *B. Boothroyd.* Protein Recovery, *M. Scawen.* Aqueous Phase Separation, *M.R. Kula.* Advances in Process Scale Chromatography, *I.B. Sabran.* Affinity Tails and Ligands, *P.D.G. Dean.* Membrane Processes, *S.J. Docksey.* Affinity Chromatography Using Monoclonal Antibodies, *C.R. Hill, J.R. Birch and C. Benton.* Continuous Electrophoresis, *C.A. Lambe and A.R. Thomson.* Separation Processes Using Supercritical Fluids, *V.Y. Loh, P. Richmond and D. Steytler.*

Special Publications of the Society for General Microbiology, No. 17

Microbes in Extreme Environments

Edited by **R.A. Herbert and G.A. Codd**
September 1986, $35.00/£19.50(UK only), ISBN: 0.12.341641.X Pbk.
September 1986, $74.00/£41.00(UK only), ISBN: 0.12.341460.1 Hbk.

Contents: Preface. The Ecology and Physiology of Psychrophilic Microorganisms, *R.A. Herbert.* The Alkaline Saline Environment, *W.D. Grant and B.J. Tindall.* Halotolerant and Halophilic Microbes, *R.H. Reed.* Fungal Responses Towards Heavy Metals, *G.M. Gadd.* Strategies of Phototrophic Bacteria in Sulphide-Containing Environments, *H. van Gemerden and R. de Wit.* Damaging Effects of Light on Microorganisms, *G.C. Whitelam and G.A. Codd.* Pressure as an Extreme Environment, *R.Y. Morita.* Bacteria Adaptations for Growth in Low Nutrient Environments, *R.J. Sharp and M.J. Munster.* Genetic Applications of Alkalophilic Microorganisms, *K. Horikoshi.* Index.

The Rat Brain in Stereotaxic Conditions

Second Edition

George Paxinos and Charles Watson
November 1986, 264pp., $53.50/£44.50(UK only),
ISBN: 0.12.547621.3

The first edition of this stereotaxic atlas has been acclaimed as outstanding for studies on the brain of the laboratory rat. It represents a modern synthesis of data on the topography of the rat's brain and provides an accurate guide to identification and stereotaxic localization of structures for both researchers and students in the neurosciences.

Physiological Ecology of Forest Production

J.J. Landsberg
August 1986, 206pp., $45.00/£25.00(UK only), ISBN: 0.12.435965.5

Recent research has produced much new information on the physiology of forest trees growing together in a stand, and on the various effects of environmental factors on them. This book provides the first summary of this subject, dealing with every aspect from micrometerology to photosynthesis, carbon utilization and water relations.

Contents: Preface. Symbols and Definitions. Introduction. Process Rates and Weather. Stand Structure and Microclimate. The Carbon Balance of Leaves. The Carbon Balance of Trees. Nutrient Dynamics and Tree Growth. Water Relations. Synthesis.

Monoclonal Antibodies: Principles and Practice

Second Edition

J.W. Goding
October 1986, c.302pp., $29.50/£17.50(UK only),
ISBN: 0.12.287021.2

The second edition of this well known and widely accepted publication is a selectively updated volume. The improvements include clarification of detail and the incorporation of the rapid advances which have been made in this field.

Contents: Preface. Preface to Second Edition. List of Abbreviations. Introduction. Theory of Monoclonal Antibodies. Production of Monoclonal Antibodies. Purification, Fragmentation and Isotopic Labelling of Monoclonal Antibodies. Analysis of Antigens Recognized by Monoclonal Antibodies. Affinity Chromatography Using Monoclonal Antibodies. Immunofluorescence. Generation of Conventional Antibodies. Glossary. Index.

CAD/CAM and MIS in Japan

Computer Applications in Japanese Industry

Edited by **Noboru Takagi, John Branch and Keiji Yajiama**
November 1986, 417pp., $79.00/£49.00 (Tentative),
ISBN: 0.12.682580.7

This is the first publication on the Ishikawa Prize lectures which reflect the efforts to modernise Japanese industrial management.

The chapters have been grouped into the following categories: (1) building and architecture, (2) services, (3) industrial management, (4) industrial production, (5) transport and construction. A historical description of the development of Japanese industry in terms of this classification, written by the economist Mr Akira Kakuta, is given in an appendix.

High Performance Liquid Chromotography

Advances and Perspectives *Volume 4*

Edited by **Csaba Horvath**
December 1986, 330pp., $70.00/£58.50(UK only),
ISBN: 0.12.312204.X

From Reviews of Volume 3:

"These contributions represent more than a survey; they are a milestone in the advancement of HPLC."

— **Analytical Chemistry**

Contents: Reversed-Phase Chromatography of Proteins and Nucleic Acids: Practical Considerations. Mixed-Interaction Chromatography. Liquid Chromatography/Electrochemistry. High-Performance Liquid Chromatography Separations of Large Molecules: A General Model.

Prices subject to change without notice

24-28 OVAL ROAD, LONDON NW1 7DX, ENGLAND • ORLANDO, FLORIDA 32887, U.S.A.

NEW AUTUMN TITLES FROM

•ACADEMIC PRESS•

A Fems Symposia Publication

Protein-Carbohydrate Interactions in Biological Systems

The Molecular Biology of Molecular Pathogenicity

Edited by **D.L. Lark, S. Normack, B.-E. Uhlin, and H. Wolf-Watz**

October 1986, 496pp., $76.50/£42.50(UK only), ISBN: 0.12.436665.1

Contents. Session I, Genetics of Bacterial Adhesins. Session II, Antigenic Variation of Bacterial Adhesins. Session III, Role of Pili and Adhesins in Pathogenicity. Session IV, Carbohydrate Receptors: Identification, Analysis and their Biological Role. Session V, Adhesins: Their Receptors and Interactions. Session VI, Bacterial Inversion. Session VII, Bacterial Surface Components and their Importance for Virulence. Session VIII, Bacterial Toxins and Receptors.

Perinatal Genetics Diagnosis and Treatment

Edited by **Ian H. Porter, Norma H. Hatcher, and Ann M. Willey**

September 1986, 328pp., $45.00/£37.50(UK only), ISBN: 0.12.562855.2

A summary of the current state of the art in perinatal genetics. It clarifies and rationalizes certain aspects of diagnosis, genetic counselling, and management of genetic disease, and serves as a useful guide for health professionals who provide care to pregnant women and their newborns.

The Mineral Nutrition of Higher Plants

Horst Marschner

October 1986, 612pp.,
$39.95/£22.50(UK only), ISBN: 0.12.473541.X Pbk.
$89.50/£49.95(UK only), ISBN: 0.12.473540.1 Hbk.

Contents: Part I: Nutritional Physiology. Introduction, Definition and Classification of Mineral Nutrients. Ion Uptake of Individual Cells and Roots; Short Distance Transport. Long Distance Transport in the Xylem and Phloem and its Regulation. Uptake of Mineral Elements by Leaves (Aerial Parts). Yield Physiology and Source-sink Relationships. Mineral Nutrition and Yield Response. Nitrogen Fixation. Functions of Mineral Nutrients — Macronutrients. Functions of Mineral Nutrients — Micronutrients. Beneficial Mineral Elements. Relationship Between Mineral Nutrition and Plant Diseases and Pests. Diagnosis of Deficiency and Toxicity of Mineral Nutrients. Part II: Soil-plant Relationship. Nutrient Availability in Soils. Root Growth and Development as Affected by External and Internal Factors. The Soil-root Interface (Rhizosphere) in Relation to Mineral Nutrition. Adaptation of Plants to Adverse Chemical Soil Conditions.

Gene Activity in Early Development

Third Edition

Eric H. Davidson

September 1986, 675pp., $49.50/£41.50(UK only), ISBN: 0.12.205161.0

This new third edition reflects the ten years of technological progress since the last edition. Providing a unique blend of classical and molecular knowledge, it discusses all major embryonic systems from both a comparative and mechanistic point of view. In deriving overall interpretations of developmental phenomena, it brings into play all the disparate forms of evidence, including genetic, molecular, and cytological.

Monoclonal Antibodies Against Bacteria

Volume III

Edited by **Alberto J.L. Macario and Everly Conway de Macario**

1986, 340pp., $52.50/£44.00(UK only), ISBN: 0.12.463003.0

Monoclonal antibodies are making a major impact on virtually every field in the health sciences as demonstrated by their increasing use for the diagnosis, prophylaxis, and treatment of diseases caused by bacteria and their toxins. This multidisciplinary treatise provides a comprehensive and up-to-date reference source on new developments of practical importance in microbiology and infectious diseases, focusing on advances made possible by monoclonal antibodies.

Military Radiobiology

Edited by **James J. Conklin and Richard I. Walker**

1986, 390pp., $34.95/£29.50(UK only), ISBN: 0.12.184050.6

All radiation biologists and biophysicians should find this an interesting and valuable reference on the effects of catastrophic ionizing radiation on the body. The book includes discussion of cellular radiation biology, physiological effects of radiation, combined injury, radioprotectants, and low-level effects.

Pattern Recognition by Humans and Machines

Edited by **Eileen C. Schwab and Howard C. Nusbaum**

Volume 1: Speech Perception, July 1986, 336pp.,
$29.95/£25.00(UK only), ISBN: 0.12.631403.9 Pbk.
$59.50/£49.00(UK only), ISBN: 0.12.631401.2 Hbk.

Volume 2: Visual Perception, July 1986, 264pp.,
$23.95/£20.00(UK only), ISBN: 0.12.631404.7 Pbk.
$46.50/£39.00(UK only), ISBN: 0.12.631402.0 Hbk.

These two volumes address fundamental problems in cognitive science, those of perceptual processing in humans and machines. They present a broad view of the field, with contributions by leading researchers in psychology, artificial intelligence, brain theory, speech and vision.

Environmental Radioactivity from Natural, Industrial and Military Sources

Third Edition

Merril Eisenbud

November 1986, 444pp., $55.00/£46.00(UK only), ISBN: 0.12.235153.3

Contents: Introduction. The Biological Basis of Radiation Protection. Radiation Protection Standards. Some General Considerations Concerning Environmental Radioactivity, Including the Use of Models. Transport Mechanisms in the Environment. Terrestrial and Aquatic Pathways. Natural Radioactivity Production and Reprocessing of Nuclear Fuels. Power Reactors. Various Other Sources of Exposure. Radioactive Waste Management. Fallout from Nuclear Explosions: I. Short-Term Effects. Fallout from Nuclear Explosions: II. Worldwide Effects. Accidents that Resulted in Contamination of the Environment. Radiation Exposure and Risk: Some Contemporary Social Aspects.

Prices subject to change without notice

24-28 OVAL ROAD, LONDON NW1 7DX, ENGLAND • ORLANDO, FLORIDA 32887, U.S.A.

Reader Service No.25

FREEMAN/
SCIENTIFIC AMERICAN

Autumn Reading from

New

Tilings and Patterns

Branko Grünbaum and G C Shephard

Visually attractive and intellectually rich, this unique, definitive book is indispensable to mathematicians, and fully accessible to any reader interested in this fascinating 'real-world' aspect of geometry.
November, 640 pages, 517 illus, board £54.95

Knotted Doughnuts and Other Mathematical Entertainments

Martin Gardner

This eleventh and latest collection of Martin Gardner's celebrated *Scientific American* columns, provides hours of enjoyment with an entirely new set of problems, paradoxes, teasers and tricks.
November, 288 pages, 207 illus, board £16.95, paper £10.95

Introduction to Comparative Cognition

Herbert L Roitblat

The first comprehensive introduction to thinking, information processing and memory in animals, with emphasis on the approaches, methods and goals of comparative cognition.
November, 384 pages, 60 illus, board £29.95, paper £17.95

Scientific American Library

Drugs and the Brain

Solomon H Snyder

A new title in the *Scientific American Library* dealing with the clinical effects of the major classes of psychoactive drugs, focusing on the use and abuse of a number of powerful and notorious drugs.
November, 256 pages, 148 illus, board £15.95

Forthcoming titles

Remote Sensing 2/e
Principles and Interpretation

Floyd F Sabins, Jr.

This acclaimed work has been thoroughly revised and updated to include explanations of the most modern remote sensing systems and the skills needed to interpret all forms of imaging technology.
December, 592 pages, 568 illus, board £47.95

Teaching Thinking Skills
Theory and Practice

Edited by Joan B Baron and Robert J Sternberg

A collection of essays by twelve eminent psychologists, educators, and philosophers that unite classical and modern theories of thought with the latest practical approaches to the learning and teaching of thinking skills.
January 1987, 304 pages, board £21.95 approx.
paper £12.95 approx.

Public Appearances, Private Realities
The Psychology of Self-Monitoring

Mark Snyder

This is the first thorough, integrated look at what the theories, research, and applications of self-monitoring tell us about our own social behavior and about the nature of the self in general.
January 1987, 320 pages, 12 illus, board £25.95 approx.
paper £13.95 approx.

W H Freeman
20 Beaumont Street, Oxford OX1 2NQ

Reader Service No.27

YALE
UNIVERSITY PRESS
13 BEDFORD SQUARE
LONDON WC1B 3JF

Origins of Sex

Lynn Margulis and Dorion Sagan

A fascinating and detailed examination of the evolution — and occasional devolution — of sexuality in microorganisms and more complex forms of life. Key issues discussed include the origin of meiotic sex through cannibalism, the connections of sex and differentiation to cell symbiosis, and the loss of sex.
Illus. £30.00

The Shaping of America

A Geographical Perspective on 500 Years of History
Volume I: Atlantic America, 1492-1800
D. W. Meinig

This entirely fresh interpretation of American history by a renowned historical geographer is the first in a projected three-volume series. In this volume, Meinig focuses on colonial America, examining how an immense diversity of ethnic and religious groups — Europeans, Africans, American Indians — ultimately created a set of distinct regional societies.
81 *b & w illus.* £35.00

The First Resource

Wild Species in the North American Economy
Christine Prescott-Allen and Robert Prescott-Allen

Christine and Robert Prescott-Allen here provide the first systematic analysis of the economic importance of wildlife to North America in a book with important implications for economic and development policies.
Illus. £58.00

Migrations in Prehistory

Inferring Population Movement from Cultural Remains
Irving Rouse

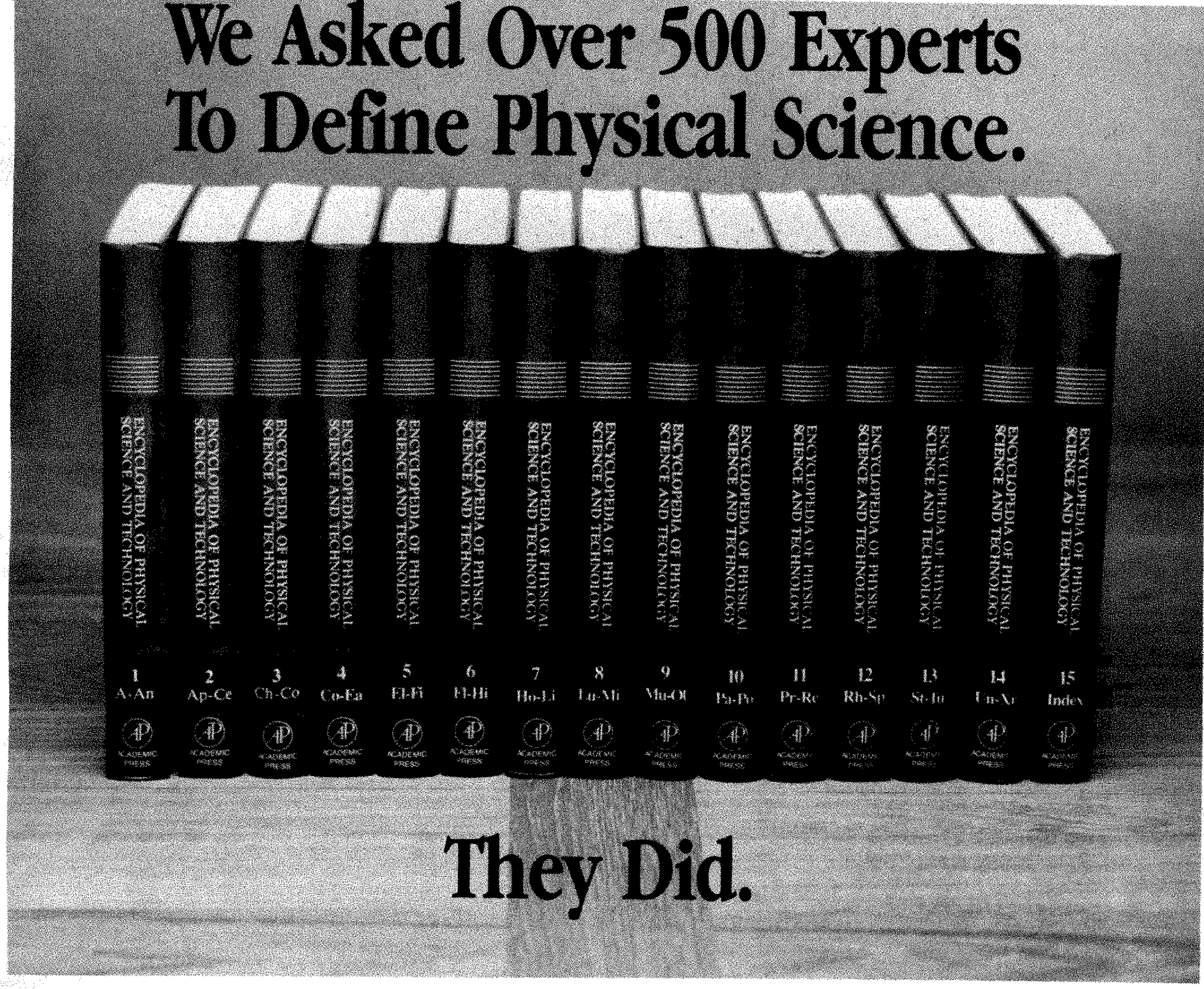
In this book Irving Rouse evaluates research on prehistoric migrations, from the successfully formed and tested hypotheses explaining the origins of the Polynesians, Eskimos, Japanese, and Tainos, to the more fanciful postulations by authors such as Thor Heyerdahl and Barry Fell. Rouse's work demonstrates the effectiveness of collaboration and communication between branches of archaeology and anthropology.
30 *b & w illus.* £20.00

The Hawk's Nest Incident

America's Worst Industrial Disaster
Martin Cherniack, M.D.
Foreword by Phillip Landrigan, M.D. and Anthony Robbins, M.D.

This engrossing and disturbing book tells the story of the worst occupational health disaster in American history, which occurred fifty years ago when hundreds of men died and over a thousand fell ill from acute silicosis contracted during the building of Union Carbide's Hawk's Nest Tunnel in West Virginia.
Illus. £19.95

Reader Service No.37



Encyclopedia of Physical Science and Technology

Robert A. Meyers, *Editor-in-Chief*

The **Encyclopedia of Physical Science and Technology** is the contemporary reference for today's high technology. It is a unique source, and the first to encompass all aspects of the physical sciences. A complete technical library, the **Encyclopedia** is:

.....DETAILED

• 15 volumes... including a separate index volume • Over 500 original articles each approximately 20 pages long • 6,000 illustrations - 2,000 tables • 11,000 double column pages • Extensive cross referencing.

.....COMPREHENSIVE

The **Encyclopedia** covers these major subjects:

• Aeronautics • Astronomy & Astrophysics • Atmospheric Sciences & Aeronomy • Chemistry • Communications • Computers & Controls • Earth Sciences & Oceanography • Electronics & Electrical Engineering • Materials Science & Engineering • Mathematics • Mechanical, Industrial, and Civil Engineering • Military Sciences • Navigation, Military Communications • Nuclear Technology • Ordnance • Physics • Power Systems, Non Nuclear & Non Propulsive • Propulsion & Fuels • Space Technology.

.....AUTHORITATIVE

All articles are written by world authorities in their fields... Nobel Prize winners and recipients of the National Medal of Science.

.....ORGANIZED

Articles are arranged alphabetically and include: table of contents, glossary of unusual terms, concise definition of subject, *in depth* presentations, extensive cross references, illustrations, tables and a bibliography.

The **Encyclopedia of Physical Science and Technology** is *the* required reference for today's scientist, engineer and student.

To request a detailed prospectus and sample article contact your usual supplier or write directly to:

Academic Press
Harcourt Brace Jovanovich, Publishers
Attn: Sales Department
Orlando, FL 32887 U.S.A.

82116—DT/IY

Reader Service No.26

CALL FOR PAPERS:

JOURNAL OF EVOLUTIONARY BIOLOGY

First Issue: January, 1988
(Will be published in October 1987)

The journal will publish original research articles in English spanning, on the one hand, *theory, botany, and zoology*, and on the other, *ecology, genetics, development, and paleontology*. It is an international journal welcoming contributions from all countries. For format, please see a recent issue of the journal Evolution.

Manuscripts may be submitted either to the
Managing Editor,
Prof. S.C. Stearns,
Zoologisches Institut,
Rheinsprung 9,
CH-4051 Basel, Switzerland,
or to a member of the Editorial Board:

Richard Abbott (St. Andrews)
Graham Bell (McGill)
William Chaloner (London)
Gary Freeman (Austin)
Tony Hallam (Birmingham)
Suresh Jayakar (Pavia)
Shoichi Kawano (Kyoto)
Eduard Linsenmair (Wurzburg)
Elisabeth Lord (Riverside)
Karl Niklas (Ithaca)
Linda Partridge (Edinburgh)
Rudolf Raff (Bloomington)
Joe Travis (Tallahassee)
Pere Alberch (Harvard)
Michael Bulmer (Oxford)
Hugh Dingle (Davis)
Pierre-Henri Gouyon (Montpellier)
Erkki Haukioja (Turku)
Gerdien de Jong (Utrecht)
Marty Lechowicz (McGill)
Volker Loeschke (Aarhus)
Hans Metz (Leiden)
Arie van Noordwijk (Basel)
Wolf Reif (Tubingen)
Valerio Sbordoni (Rome)
Gunter Wagner (Vienna)

First Annual Meeting of the EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY 26–30 August 1987 Basel, Switzerland

Plenary lectures by distinguished speakers. Workshops and symposia with contributed talks.

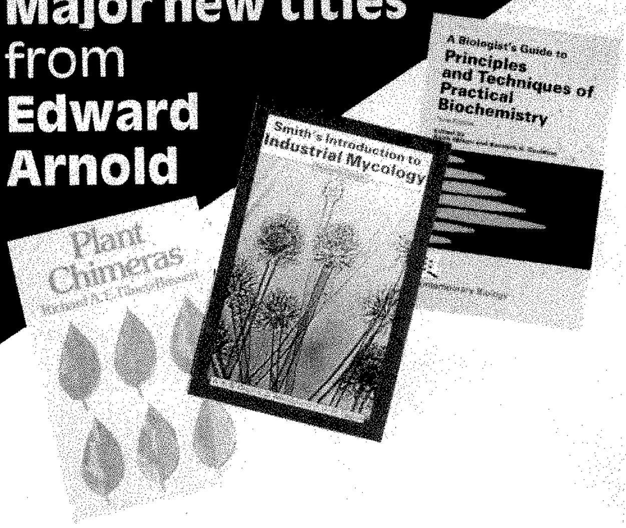
A business meeting to approve bylaws, elect officers, and establish the Society's relationship to the Journal of Evolutionary Biology.

Please address suggestions and requests for information to: Dr. A.J. van Noordwijk, Zoologisches Institut, Rheinsprung 9, CH-4051 Basel, Switzerland.

Birkhäuser Verlag
Basel · Boston · Stuttgart

Reader Service No.61

Major new titles from Edward Arnold


A Biologist's Guide to Principles and Techniques of Practical Biochemistry
Smith's Introduction to Industrial Mycology
Plant Chimeras
Richard A.E. Tilney-Bassett

Plant Chimeras

R.A.E. Tilney-Bassett

This book introduces the main steps in the analysis of chimeras, explains their structure and development, and the ways of classifying and manipulating them. Each chapter is illustrated by examples taken from the extensive international literature, show the many and varied facets of chimera growth and behaviour and the interactions between component tissues.

Publication September 1986
208 pages 35 line and 22 half-tone illustrations
£13.50 net paper 0 7131 2936 0

A Biologist's Guide to

Principles and Techniques of Practical Biochemistry

THIRD EDITION
Edited by **K. Wilson** and **K.H. Goulding**

'. . . continues to be an excellent, clearly written textbook.' *Biologist* on the last edition. This book concentrates on those up to date techniques commonly used in biochemistry, which all students of biology would be expected to encounter during their course of study. All chapters have been revised and two new chapters cover enzyme assay techniques and methods used in molecular biology.

Publication October 1986
416 pages line and half-tone illustrations
£14.50 net paper 0 7131 2942 5

Now in paperback . . .

Smith's Introduction to Industrial Mycology

SEVENTH EDITION
A.H.S. Onions, D. Allsopp and **H.O.W. Eggins**

The intention of this book is to produce a modern successor to G. Smith's *Introduction to Industrial Mycology*. '. . . a clearly written and very well illustrated textbook on mycology with special emphasis on taxa relevant to industrial mycology . . . recommended to every serious student of mycology.' **International Journal of Mycology and Lichenology**

Probable publication November 1986
408 pages 85 line and 142 half-tone illustrations
£19.95 net approx paper 0 7131 2949 2
(originally published in hardback 1981)

Edward Arnold
41 Bedford Square, London WC1B 3DQ

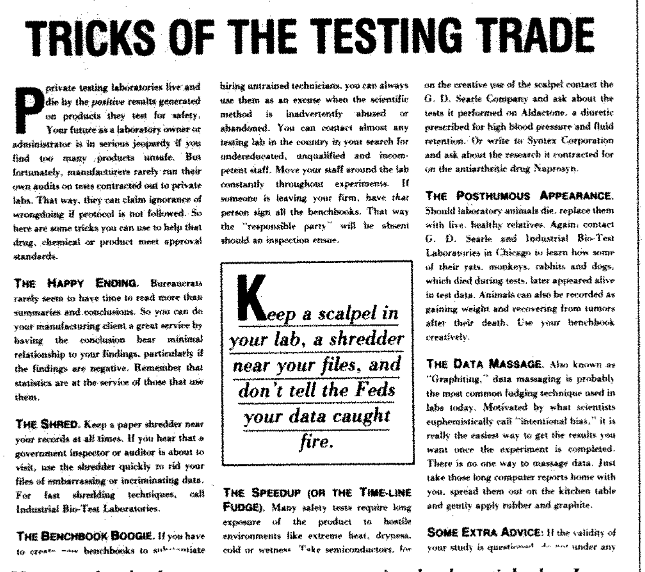
Reader Service No.36

Deception beyond measure

C.R.B. Joyce

False Prophets: Fraud and Error in Science and Medicine. By Alexander Kohn. *Basil Blackwell: 1986. Pp.221. £14.95. To be published in the United States in January 1987, $24.95.*

As THIS book and its chief rival, William Broad and Nicholas Wade's *Betrayers of the Truth** both remind us, Charles Babbage distinguished three unorthodox ways of dealing with data. These he called, in increasing order of impropriety, trimming, cooking and forging. Are such behaviours more common today than they were in Babbage's time? In his very first sentence Alexander Kohn states that the exponential growth of science has been accompanied by the "appearance" of deviant scientists, as if a rectilinear growth would not have given rise to the problem. It is not clear whether he thinks that there has been an increase in the number of such individuals, as Broad and Wade imply; and, if so, whether this is absolute or proportional. In fact, there have been more attempts than Kohn realizes to answer the question, but it is not an easy one to tackle, requiring as it does methods learnt from the confessional, the KGB and Social Audit. But whether or not the wickedness has increased or is increasing, we are sure that it should be diminished.



TRICKS OF THE TESTING TRADE

Private testing laboratories live and die by the *positive* results generated on products they test for safety. Your future as a laboratory owner or administrator is in serious jeopardy if you find too many products unsafe. But fortunately, manufacturers rarely run their own audits on tests contracted out to private labs. That way, they can claim ignorance of wrongdoing if protocol is not followed. So here are some tricks you can use to help that drug, chemical or product meet approval standards.

THE HAPPY ENDING. Bureaucrats rarely seem to have time to read more than summaries and conclusions. So you can do your manufacturing client a great service by having the conclusion bear minimal relationship to your findings, particularly if the findings are negative. Remember that statistics are at the service of those that use them.

THE SHRED. Keep a paper shredder near your records at all times. If you hear that a government inspector or auditor is about to visit, use the shredder quickly to rid your files of embarrassing or incriminating data. For fast shredding techniques, call Industrial Bio-Test Laboratories.

THE BENCHBOOK BOOGIE. If you have to create new benchbooks to substantiate

hiring untrained technicians, you can always use them as an excuse when the scientific method is inadvertently abused or abandoned. You can contact almost any testing lab in the country in your search for undereducated, unqualified and incompetent staff. Move your staff around the lab constantly throughout experiments. If someone is leaving your firm, have that person sign all the benchbooks. That way the "responsible party" will be absent should an inspection ensue.

Keep a scalpel in your lab, a shredder near your files, and don't tell the Feds your data caught fire.

THE SPEEDUP (OR THE TIME-LINE FUDGE). Many safety tests require long exposure of the product to hostile environments like extreme heat, dryness, cold or wetness. Take semiconductors, for

on the creative use of the scalpel contact the G. D. Searle Company and ask about the tests it performed on Aldactone, a diuretic prescribed for high blood pressure and fluid retention. Or write to Syntex Corporation and ask about the research it contracted for on the antiarthritic drug Naprosyn.

THE POSTHUMOUS APPEARANCE. Should laboratory animals die, replace them with live, healthy relatives. Again, contact G. D. Searle and Industrial Bio-Test Laboratories in Chicago to learn how some of their rats, monkeys, rabbits and dogs, which died during tests, later appeared alive in test data. Animals can also be recorded as gaining weight and recovering from tumors after their death. Use your benchbook creatively.

THE DATA MASSAGE. Also known as "Graphiting," data massaging is probably the most common fudging technique used in labs today. Motivated by what scientists euphemistically call "intentional bias," it is really the easiest way to get the results you want once the experiment is completed. There is no one way to massage data. Just take those long computer reports home with you, spread them out on the kitchen table and gently apply rubber and graphite.

SOME EXTRA ADVICE: If the validity of your study is questioned, do not under any

How to buck the system — a tongue-in-cheek article by Irene Moosen and Mark Dowie which appeared in Mother Jones *in 1982. (Reproduced from* False Prophets.)

When a question arises in the science of science (a respectable discipline, the methods of which have been successfully applied to the growth of science as a whole), one would expect scientists to try to answer it scientifically. Too often, however, they fall back on dogmatic or wistful assertions of belief. If public attitudes to science were still as starry-eyed as they were even 25 years ago, such piety might suffice. But reassurances that there is nothing to worry about — because the scientific process is self-corrective, and that mistakes, deliberate or accidental, will be found out sooner rather than later — are no longer enough. Such reasoning belongs to the "tap-dancing and dinosaurs" class ("I'm tap-dancing to keep the dinosaurs away!". "But there haven't been any dinosaurs for years!". "There! You see?"). That the only examples of fraud that have been discovered are the ones that have been discovered is not likely to convince the critics of science, and should not satisfy scientists either. Even Popper considers only that errors are "fairly often, in time, corrected".

How much time is needed for the purpose? Seventeen years or so between deed and discovery in the case of Cyril Burt, or 40 for the Piltdown hoax or Gregor Mendel? There are even more important questions to ask about the ethics of scientific discovery than this, but they are seldom examined systematically. One obvious framework is: What is scientific fraud (definitions)? How much of it is there (estimates, including trends)? Why does it occur (reasons — social as well as individual)? Does it matter anyway (costs — and possible benefits: see below)? And what is being and should (or can) be done about it?

Of course, Kohn is aware of these problems but, as is common among popularizers of science, he seems to believe that readers prefer a stream of hair-raising stories to a serious discussion of ways to identify and prevent fraud in the future. Thus it would serve little purpose to quote the table of contents; the divisions between subject areas often seem to be arbitrary and the chapter titles usually promise more than the text delivers. Even though his final chapter is entitled "What Can Be Done about Fraud in Science?", and might have been expected to bring together and add to suggestions scattered about the book, it is only three pages long and one of them is taken up by introducing two further anecdotes warning would-be whistle-blowers that theirs is a very chancy and perhaps even dangerous activity.

The most relevant point is made in the very last paragraph of all: "To understand deceit in science, one has to study first the ethics of the society within which the scientists work and live . . .". A useful book on scientific fraud would begin, instead of ending, at that point. Given one society (Britain) in which the world's largest advertising agency can switch from helping a national health campaign to promoting the very product whose use it had previously sought to diminish; another (the United States) in which 25,000 unnecessary cardiac pacemakers are inserted per year; a third (the Soviet Union) in which. . . . The list can be continued to infinity.

One should not, of course, criticize an author for writing the book he wanted to write; but it is fair to suggest that Kohn has not carried out his own expressed intention. His sub-title is *Fraud and Error in Science and Medicine*, less alarmist than Broad and Wade's *Fraud and Deceit in the Halls of Science* more open-minded too. It also allows Kohn to mention Mendel and Newton more properly in his book than Broad and Wade in theirs. But though with these references, as well as elsewhere (for example, in the attractive dust-jacket of Wyck's "Alchemist") Kohn implies that error may sometimes be only an idea whose time has not yet come, the theme nowhere receives the discussion its importance requires and the title of the book promises: When is an error not an error but a flash of inspiration? Not only do our hypotheses depend upon our perceptions, but perception is itself a prisoner of hypothesis. Even good scientists have sometimes been so sure that they were right that, in Alan Franklin's words, they "manufactured data . . . until real data became available" — even one of Darsee's retracted papers has already been de-retracted!

Perhaps Kohn fails to deliver because, understandably, he is as uncertain about the matter as the rest of us. In fact, he is at times so uncertain in his handling of certainty in general — for example, the meaning of correlation, and the relationship of truth to the laws of chance; not to speak of his confusion between experimenter (or cognitive) error and experimental (or statistical) error — that the reader is sometimes driven to doubt his grasp of the matters of which he writes.

According to Stephen Toulmin, "medicine has saved the life of ethics" (*Perspectives in Biology and Medicine* **25**, 736–750; 1982), producing "spectacular and irreversible effects on the methods and content of philosophical ethics". It is high time for at least a first instalment towards repayment of the debt. ☐

C.R.B. Joyce is Head of Project Innovation in the Medical Department of Ciba-Geigy Ltd, CH-4002 Basel, Switzerland.

---

*Published in 1983 by Simon & Schuster, and reviewed in *Nature* **302**, 774 (1983).

The chips are now down?

Steve Woolgar

Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. By Hubert L. Dreyfus and Stuart E. Dreyfus. *Basil Blackwell/ Free Press: 1986. Pp.231. £15, $16.85.*

THE most important and exciting thing about artificial intelligence (AI) is that it provides the possibility of a critical test of fundamental assumptions about the nature of knowledge and reason. These assumptions have gripped Western thought for 2,000 years and they underpin ingrained ideas about the character of man, his abilities and potential. If the project to devise artificial reason is adjudged a success, it will vindicate the objectivist and rationalist philosophies which are thought to have guided the rise of science and technology; its failure will buttress those phenomenological and intrepretivist philosophies which argue that calculation and measurement are largely irrelevant to man's capacity for intelligent behaviour.

Hubert Dreyfus recognizes what is at stake. His *What Computers Can't Do*, first published in 1972 and revised in 1979, was a compellingly sophisticated analysis of the philosophical assumptions which underlie AI. That book provided an overview of developments in AI from work on cognitive simulation in the 1950s to early attempts at knowledge engineering in the 1970s. *Mind Over Machine*, written with his brother (a computer specialist), extends the analysis to recent work, particularly on expert systems. The book's appearance is timely, in view of the sudden and massive increase in financial support for AI, especially from military sources.

Will these people collect on their expensive bet? The Dreyfuses say no. Their central thesis is that intelligent human behaviour (interpretation, understanding, recognition, puzzle solving, rule application and so on) is simply not amenable to formalization in the way required by digital computers.

The authors propose a model comprising five stages of skill acquisition, ranging from novice to expert. Progression beyond the third stage ("competence") requires a qualitatively different approach to reasoning. Expertise, in their view, involves the accumulation of experience, the unconscious holistic recognition of the familiarity of new situations and the critical use of intuition and "know-how". This is beyond the ability of systems designed on principles of calculation from discrete facts and codified rules. The apparent success enjoyed by some artificial devices occurs in restricted (artificial) areas which require competence rather than intelligence. This is not to deny the usefulness of some AI systems as aids, but such systems are properly termed "competent" rather than "expert". The five-stage skill hierarchy is used to discern the limits of computer use both in education and in managerial decision-making. Only in the rare instances, where researchers abandon the mechanistic models of human skill, do we find hope for a better approximation to human expertise.

The anti-mechanistic view of human skill derives from the Wittgensteinian argument that the determination of interpretation, understanding, recognition and so on is entirely construed in a social context and can never sensibly occur in isolation from it. But the Dreyfuses' hierarchy of skills is strikingly individualistic and mentalistic. For example, they characterize expertise in terms of inner processes: "the expert driver becomes one with his car, and he experiences himself simply as driving, rather than as driving a car . . ." (p.30). Given the philosophical orientation of their critique, there is surprisingly little acknowledgement of the way in which such experiences are meaningful only as socially sanctioned and publicly accountable descriptions. Although the authors insist that their aim is not to mystify, their treatment of skills as inner processes and experiences provides too easy a target for the rationalists. Similarly, it is surely offering too much of a hostage to fortune to continue to speak of the brain and the mind as receptacles of intelligence. The Dreyfuses intend these curious and outdated concepts as a sanctuary for unformalizable human skills. But as others such as Jeff Coulter have shown, notions of "brain" and "mind" are social constructions which tend to support the mechanistic view of human behaviour.

"IT FIGURES. IF THERE'S ARTIFICIAL INTELLIGENCE, THERE'S BOUND TO BE SOME ARTIFICIAL STUPIDITY."

Cartoon by Sidney Harris

The Dreyfuses are especially impressive in demonstrating that despite changes in emphasis and direction (for example, the shift from attempts to fashion general, problem-solving devices to those able to make inferences within a highly specialized knowledge domain), workers in AI repeatedly face the same fundamental conceptual problems. For the Dreyfuses, AI's failures are the direct result of such irresolvable problems. For many AI enthusiasts, however, these are merely temporary, technical and surmountable difficulties. For example, whereas the Dreyfuses say that work on "expert" systems shows it is wrong to assume that human experts are using heuristic rules, proponents of AI interpret the research as merely indicating some snags in eliciting these rules. (It is notable that most previous responses to Hubert Dreyfus's ideas have preferred to contest the alleged failure of particular systems, rather than address the conceptual problems he raises.)

Early restricted successes gave rise to unwarranted optimism about the future of AI and, for the Dreyfuses, this optimism partly accounts for the current inability of the AI community to recognize failure. But the Dreyfuses also suggest that the stakes are now so high that failure is actively concealed. They thus allege deliberate cover-up by the AI establishment, describing in the book their own treatment and the manipulation of TV appearances by the "artificial intelligentsia", and criticizing the "unscientific" attitude of AI workers. (The dust-jacket speaks of "AI's record of barefaced public deception".)

This recourse to a kind of conspiracy theory is something of a distraction. For a full appreciation of the dynamics of the AI enterprise, we need to look carefully at the social attribution of "success" and "failure" in the context of a pervasive commitment to rationalistic models of intelligence. Although the Dreyfuses' conceptual argument provides a good, sceptical springboard from which to launch a social analysis of AI research, we should avoid their supposition that conceptual problems will necessarily give rise to "failures".

Who, in any case, are those who cover up failures and spread false optimism about AI? It is important to be clear that the Dreyfuses identify a relatively small body of researchers. True, these are the most influential and respected champions of AI and are thus a legitimate and necessary target for attack; many are entrepreneurs with a substantial interest in furthering public support for the enterprise. But it might be misleading to accept their proclamations as characteristic of AI as a whole. Searle's distinction between "strong" and "weak" AI is useful here. Whereas proponents of strong AI (who claim, for example, that programs can "understand") are appropriate targets for the Dreyfuses, the more cautious advocates of weak AI (who see computers as, at best, a tool for aiding our investigations into the nature of intelligence) might agree that it is unreasonable to expect the success of artificial systems beyond the level of competence.

Weak and strong AI are not unrelated, of course. Proponents of the former have undoubtedly benefited from the public persuasiveness of the latter. But the distinction points to a further shortcoming in the Dreyfuses' case. By concentrating upon proponents of strong AI, the arguments give us little insight into what practitioners of AI do. In this book, the Dreyfuses continue their impressive assault in the titanic struggle between phenomenological philosophy and the rationalist assumptions of AI's public guardians. But we are not told whether or how the assumptions of strong AI touch upon the day-to-day business of devising systems, making programs work, evaluating output and so on.

It is interesting to consider the likely fate of the Dreyfuses' thesis. Their case against AI is the strongest and most eloquent yet published, a protest against conceptions of the human being which have held sway since the beginnings of Western thought, and it suggests that one outcome of the AI debate might be a definitive refutation of the rationalist metaphysic. However, the lack of a fully social conception of the practice of machine intelligence is perhaps reflected in their inattention to one of the key lessons of the history of science: "critical tests" are rarely decisive. Evaluations of "success" and "failure" need only generate further redefinitions of AI, not its abandonment. ☐

Steve Woolgar is a Lecturer in the Department of Human Sciences, Brunel, The University of West London, Uxbridge, Middlesex UB8 3PH, UK. With Bruno Latour he is author of Laboratory Life: The Construction of Scientific Facts, *a revised edition of which has recently been published by Princeton University Press.*

Wondrous gifts

Peter Bryant

Nature's Gambit: Child Prodigies and the Development of Human Potential. By David Henry Feldman with Lynn T. Goldsmith. *Basic Books:1986. Pp.259. $19.95.*

EVERY now and then a child turns out to be outstandingly talented in a quite specific way. It is not just that he quickly outstrips all his contemporaries in a particular skill: he also soon becomes a great deal better than most grown-ups at it. Children such as this arouse great interest and wonder. Psychologists in particular have to pay attention to them because at first sight they seem to break most of the rules set out by developmental theories. After all, these theories generally hold that different intellectual skills are connected and that they develop together at roughly the same pace through childhood. How can such theories possibly explain a five-year-old child whose mathematical skills already outstrip those of 95 per cent of the adult population, but who otherwise seems pretty ordinary?

I suspect that this question was the starting point for David Feldman's book, which is a prolonged and detailed study of six different, prodigiously talented American boys and their families, but it is not in the end a question that he answers satisfactorily. Instead he spends most time on another question which is unanswerable, or at any rate cannot be answered from a study of individual child prodigies. His main concern is to find what conditions are needed to produce such children. Feldman's own idea, which is about what he calls the "coincidence process" is neither new nor sophisticated: it is that these children have extraordinary gifts ("proclivities" is Feldman's term) to start with but that their talents only flourish in the right environment. Prodigies need sensitive and devoted parents and an appropriate social environment, according to Feldman, and he is plainly impressed by the lengths to which some of the parents in his little group were prepared to go to support their children. Sacrificing jobs, commuting long distances, confronting awkward teachers, taking the child from one tutor to the next — these seem to be typical demands for a prodigy's parents, and in this group at least they shouldered the burden quite cheerfully. Often the parents' own talents matched their child's: the boy with an extraordinary gift for music had a musician for a father, and the father of the boy who showed early signs of being a talented writer was a writer himself. In these cases the parents spent a great deal of time teaching their own specific skill to their child.

Feldman's argument that all the children in his group were lucky enough to be born into a suitable environment is convincing, but it does not prove his case that such environments are essential for the emergence of a prodigy. The trouble is that his thesis is just as much about children whom he did not study as about the six prodigies with whom he spent so much time. There must, he argues time and again, be children who could have been prodigies, who had the "proclivities", but who found themselves in the wrong environment: the world must be littered with mute, inglorious Miltons or Cromwells guiltless of their country's blood. But Feldman did not study any such prodigy *manqué*. He does not even demonstrate that they exist, and it is very difficult to see how he could have done so, given his methods.

Making music — the prodigiously talented Mozart, accompanied by his father and sister.

Feldman's description of the lives of the six boys is interesting and frequently poignant: but it is difficult to see what is new here. After all, we already know a great deal about the childhood of erstwhile prodigies. Mozart's childhood, to take one example, is documented as well as any other part of his life. Sartre, to take another, wrote about his early years in great detail and with considerable distaste. One of the most striking features of these examples and the children studied by Feldman is the surprise and interest that the prodigies excited in the adults around them. The children soon become performers and they do so at the insistence of adults. On occasion, adult astonishment at the child's unexpected talents has some odd results. In my view the most interesting part of Feldman's book is his description of the way in which parents or teachers sometimes resort to irrational explanations for a child's extraordinary talents, because they cannot account for them in any other way. Reincarnation is one solution. The child is thought to be so talented because he is simply a new version of a dead genius.

To read Feldman's sympathetic and sensitive accounts of these six children is to feel the same kind of awe which the young Mozart must have inspired in those who heard him play and improvise so effortlessly. Readers of this book will wonder at the phenomena which Feldman describes, but will remain puzzled about the reasons for them and about the implications for developmental psychology. ☐

Peter Bryant is Watts Professor in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

Knockabout stuff

Derek Ager

The Tenth Muse: The Pursuit of Earth Science. By Ronald B. Parker. *Charles Scribner's Sons: 1986. Pp. 221. $15.95.*

THE nine official muses were responsible for literature, art and a limited amount of science. They had originally been three in number, but six more were conceived and so it is reasonable to add another to represent geology, the most beautiful and inspiring of the sciences. Professor Parker, however, obviously nominates the subject for its entertainment value rather than for its aesthetic qualities.

We are all interested chiefly in our own ideas, but Professor Parker's egocentricity is particularly evident. He assumes that we have read his previous book, and tells us how his great "buffer" idea came to him and how the audience reacted to it at various lectures. Personally I do not understand the "buffer" terminology which is central to most of Dr Parker's thoughts, and would have thought that a better metaphor for what he discusses (the storage and sudden release of energy) would have been the elastic band, which can be stretched so far before suddenly breaking.

Parker's main interest is in our geological today and yesterday — geysers and landslips, glaciers and ice dams. In "The Urge to Surge", for example, he discusses the movement of glaciers and their episodic behaviour, making an assignation with the reader up in Alaska around the year 2000 when things are due to get moving again.

In an excursion into his other main field, crystallography, Parker points out that this science originated in the West, where symmetry is taken for granted, and not in the East where the Zen philosophy prefers asymmetry, be it in gardens or pottery. He goes on to discuss man and other "non-crystallographic" forms in the world (with a delightful independence of thought, he and his wife built themselves a 14-sided house), and Schechtman's discovery of the impossible, that is to say, five-fold symmetry in crystals, as a salutary warning of the need for his own kind of open-mindedness. As a palaeontologist I should point out that echinoderms came up with five-fold symmetry first, albeit using trigonal calcite crystals.

There is a boyhood trip to Boron, site of the largest borate mine in the world, before an essay on "coral clocks" and organic growth generally. Here Parker regrets that corals, with their daily growth-lines, are rare as fossils compared with "clams", thereby displaying a surprising ignorance of the Palaeozoic part of the record. He obviously prefers the "clams" (including their suitability for feeding the poor graduate students studying them), and shocks the palaeontologist by referring to the brachiopods as their "look-alikes". All of this makes me prefer Parker as a geomorphologist and crystallographer rather than as a palaeontologist–stratigrapher.

So we go on to "Aqua Spectaculars and Other Disasters", ranging from Esther Williams and her symmetrical swimmers in those boring films of the 1930s, and via an extremely tenuous link to tsunamis and catastrophically emptying lakes, such as Lake Bonneville and, of course, J. Harlen Bretz's spectacular Lake Missoula. And so back to the "buffer" phenomenon (or better perhaps the knicker elastic of the same vintage as Professor Parker's films) in the building and release of energy in eruptions ranging from Mount St Helen's to Santorini.

There is, too, a general discussion of fossil "scat" (excreta to British readers), including description of a remarkable concentration of three billion tons of apatite in the Kola Peninsula which *might* have been the result of a monstrous ball of dung hitting the Earth. More seriously, we are told that from the evidence of their "scat" the giant ground sloths may have become extinct from a drug overdose, through eating the plant known as "Mormon tea". So the book goes on, leaping nimbly from one unlikely topic to another — remote sensing to salt "punch-ups", karst erosion to mass extinctions, until finally reaching the geological glory of field-work in (where else for a muse?) the Elysian Fields.

This is an enjoyable book, a cheerful somersault rather than a solemn symposium and better for entertainment than for education. But that, presumably, was the author's intention. □

Derek Ager is Research Professor in the Department of Geology, University College of Swansea, Swansea SA2 8PP, UK.

Wilder shores

Peter J. Smith

Islands. By H. W. Menard. *W. H. Freeman: 1986. Pp.230. £15.95. In the United States available only to members of the Scientific American Library.*

THERE seem to be as many books called *Islands* as there are islands, concentrating on everything from scenery, through seafaring and survival, to science. Menard's version falls into the last category, as one might expect from a distinguished geological oceanographer who ended up as Director of the US Geological Survey. But a scientific book on islands sounds a bit like one on kidneys; one can see the attraction of specialization whilst taking leave to doubt whether much sense can be made of the parts in isolation from the body as a whole. As far as the Earth sciences are concerned, the revolution of the 1960s, which produced the first convincing global theory of the Earth's behaviour, placed islands firmly in the context of spreading sea-floors and jostling tectonic plates, from which conjunction they are never likely to escape.

Inevitably, therefore, Menard has ended up writing a book largely about plate tectonic processes, albeit with islands taking centre stage and with little mention of continents. Excluded too, incidentally, are those islands, large or small, with a continental structure (Britain, New Zealand, Japan, the Seychelles, certain Indonesian islands and so on). Menard's islands are oceanic — volcanic or volcanic/coral, "small, young, isolated, simple, and subjected to a limited range of environmental factors". In geological terms they represent the closest that nature comes to laboratory conditions, with dependent variables at a minimum.

Notwithstanding the restricted definition, there is a surprising amount to say. For one thing, islands are perpetually under the influence of spreading lithosphere and stationary hot spots, and all that means in terms of aging, subsidence, rates of extrusion, distribution, redistribution and, ultimately, subduction, continental underplating or terrane formation. New volcanic masses lead to isostatic adjustments and lithospheric warping, the effects of which can be obscured by multicausal changes in sea level. Then there is the whole phenomenology of underwater volcano formation and decay, involving magma flow, growth rates, erosion and submergence, with or without the participation of coral. Finally, the issue of colonization raises important questions about evolution, migration and civilization.

Plate tectonics though much of this may be in essence, it is plate tectonics viewed obliquely. It does no harm at all to be forced to look at familiar concepts from an unusual angle as long as the viewpoint is coherent, which is Menard's hands it certainly is. The level could best be described as *Scientific American*, which means that the absolute novice may find it tough going in parts. Those with a smattering of basic terminology, however, will discover a rewarding and unusually enjoyable volume, not at all a bad memorial to an Earth scientist who was ever ready to look beyond his immediate research. □

Peter J. Smith is Reader in Earth Sciences at the Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

- Another of Henry Menard's books to be published posthumously, *The Ocean of Truth: A Personal History of Global Tectonics* (Princeton University Press), will be reviewed in a future issue of *Nature*.

New Books from IRL Press

THE PRACTICAL APPROACH SERIES

Carbohydrate Analysis

Edited by M F Chaplin and J F Kennedy

Although there is a vast literature covering numerous carbohydrate analytical methods, it is not always obvious to the researcher which method, or even physical technique, is the most appropriate to a particular investigation. *Carbohydrate Analysis* answers the need for a handbook of laboratory protocols in this field. It brings together the proven techniques, compares their merits in different situations, and then provides full practical details of the most suitable methods within those techniques.

Due November 1986; approx 240 pp;
Soft 0 947946 44 6; £16/US$29
Hard 0 947946 68 3; £26/US$47

Spectrophotometry and Spectrofluorimetry

Edited by D A Harris and C L Bashford

Biochemists who use this new book will see clearly how spectrophotometry can contribute to laboratory analyses – what is required in the experimental set up and how the information obtainable should be optimized. Researchers and clinicians will find the detailed protocols and thorough introduction to the use of fluorimeters and spectrophotometers invaluable, particularly if they are using them for the first time or with only a rudimentary knowledge. Technicians carrying out routine analyses will also benefit from this laboratory handbook.

Due November 1986; approx 186pp;
Soft 0 947946 46 2; £16/US$29
Hard 0 947946 69 1; £26/US$47

Biochemical Toxicology

Edited by K Snell and B Mullock

Describes the most useful and recent procedures available for biochemical toxicology. Particular emphasis is given to the application of general biochemical techniques to toxicological research. Experienced toxicologists offer not only details of protocols, but also their all-important 'tricks of the trade', which are often omitted from research papers.

Due December 1986; approx 256pp;
Soft 0 947946 52 7; £16/US$29
Hard 0 947946 67 5; £26/US$47

Nucleic Acid and Protein Sequence Analysis

Edited by M Bishop and C Rawlings

The determination of nucleic acid and protein sequences is a technique fundamental to molecular biology and biotechnology. This new book describes practical methods using computer analysis in the following areas: experimental determination of restriction maps and DNA sequences, predicting DNA function, predicting RNA and protein folding, searching databases for homologous sequences and making inferences about molecular evolution.

Due December 1986; approx 360pp;
Soft 1 85221 006 0; £16/US$29
Hard 1 85221 007 9; £26/US$47

Pre-Publication Prices

Prices quoted are valid for all pre-paid orders received before 31 December 1986. Thereafter prices may vary.

COMPLETE CATALOGUE AVAILABLE ON REQUEST

US$ prices valid in North and South America and Japan, £ prices elsewhere. All orders sent direct to publisher should be prepaid – publisher pays postage. Orders received without payment will be first invoiced to customer and postage added. Payment must be received before books are despatched.

ORDERS TO

IRL Press Ltd,
PO Box 1, Eynsham,
Oxford OX8 1JJ, UK

IRL Press Inc,
PO Box Q, McLean,
VA 22101-0850, USA

Plasmids

Edited by K G Hardy

Provides all the essential methods for studying bacterial plasmids with emphasis on the most up-to-date and proven methods. Written by acknowledged experts, the chapters have been chosen for their central importance or because they provide unique descriptions of methods for areas of study currently undergoing rapid development. An essential source book for advanced courses in microbial genetics and for laboratory research.

Due December 1986; approx 240 pp;
Soft 0 947946 81 0; £16/US$29
Hard 0 947946 82 9; £26/US$47

Teratocarcinomas and Embryonic Stem Cells

Edited by E J Robertson

Embryonic stem cell cultures are now widely used as an experimental model system by investigators sharing diverse primary scientific and clinical interests. This new book describes both the basic, routine techniques and the specialist techniques associated with this experimental system. Provides in-depth coverage of the practical aspects of cultivated stem cell lines, include the more recently developed techniques for isolation of murine pluripotent stem cells (ES or EK cells).

Due December 1986; approx 256pp;
Soft 1 85221 004 4; £16/US$29
Hard 1 85221 005 2; £26/US$47

ALSO OF INTEREST

Molecular Medicine Volume 2

Edited by A D B Malcolm

Applications to clinical practice of recent advances in tissue and organ biochemistry.

Contents:
Hyperthermic therapy and heat shock response *R H Burdon* • The molecular genetics of lipid metabolism *J Scott* • Blood coagulation and its abnormalities *I Peake and M Liddell* • Vaccine design: A challenge to molecular biology *M Mackett.*

Due December 1986; approx 180pp;
Soft 0 947946 58 6; £16/US$29

Biological Monitoring of the State of the Environment: Bioindicators

IUBS monograph series no 1
Edited by J Salanki
Due October 1986; 80pp;
Soft 1 85221 015 X; £12/US$20

Biological Nomenclature Today

IUBS monograph series no 2; papers presented at ICSEB III; held in Brighton, UK; 9 July 1985.
Edited by W D L Ride and T Younes.
Due October 1986; 80pp;
Soft 1 85221 016 8; £12/US$20

Nucleic Acids Symposium Series No 17

Proceedings of the Fourteenth Symposium on Nucleic Acids Chemistry; Japan 1986.
Due October 1986; 238 pp;
Soft 1 85221 003 6; £27/US$49

IRL PRESS

Reader Service No.41

Autumn books

Fundamentals of Immunology

By **O. G. Bier, W. Dias da Silva, D. Götze, I. Mota**

With a contribution by R. Burger

2nd, completely revised edition. 1986. 173 figures. Approx. 480 pages. Soft cover DM 64,–. ISBN 3-540-15332-2

"The book's stated aim is to provide an introduction, primarily for medical and biology students, as a foundation for subsequent reading. It does this certainly, and with a clarity, sense of purpose and economy that is not achieved in many other recent texts." *immunology today*

F. Vogel, Heidelberg; **A. G. Motulsky**, Seattle

Human Genetics

Problems and Approaches

2nd, completely revised edition. 1986. 447 figures, 210 tables. XXXIV, 807 pages. Hard cover DM 148,–. ISBN 3-540-16411-1

"This book should be read by all who work in human genetics and who teach it to others. It should be read by all advanced students in genetics, human or otherwise, and it should be read – or at least the nontechnical parts of it – by those outside the field who would like insight into the way human genetics and human geneticists work…" *American Journal of Human Genetics*

J. C. Rüegg, University of Heidelberg

Calcium in Muscle Activation

A Comparative Approach

1986. Approx. 97 figures. Approx. 320 pages. (Zoophysiology, Volume 19). Hard cover DM 198,–. ISBN 3-540-17117-7

This monograph presents a unifying picture of the control of calcium movements and calcium action in smooth, cardiac, and skeletal muscle.

E. Gwinner, Andechs

Circannual Rhythms

Endogenous Annual Clocks in the Organization of Seasonal Processes

1986. 73 figures. XI, 154 pages. (Zoophysiology, Volume 18). Hard cover DM 128,–. ISBN 3-540-16891-5

Contents: Introduction. – Evidence for Circannual Rhythms. – Properties of Free-Running Circannual Rhythms. – Synchronization of Circannual Rhythms. – Mechanisms of Circannual Organization. – Adaptive Significance of Circannual Rhythms. – Appendix: General Oscillator Model and Terminology. – References. – Systematic Index. – Subject Index.

E. Koushanpour, Northwestern University; **W. Kriz**, University of Heidelberg

Renal Physiology

Principles, Structure and Function

2nd edition. 1986. 193 figures, 3 color plates. XII, 390 pages. Hard cover DM 124,–. ISBN 3-540-96304-9

This second revised edition presents an integrated and quantitative analysis and synthesis of the subjects of fluid and electrolytes, kidney structure and function, acid-base balance, and their regulation under normal conditions. The purpose of this book is to provide the latest comprehensive information about the above subjects for those interested in the general field of nephrology.

Small-Bowel Transplantation

Experimental and Clinical Fundamentals

Editors: E. Deltz, A. Thiede, H. Hamelmann, Kiel

1986. Approx. 400 pages. Hard cover DM 180,–
ISBN 3-540-16336-0

For the first time, the results of all groups of scientists working in the new field of small bowel transplantation are compiled in one volume. The book provides fundamentals, including animal models, immunology, and immunosuppression, as well as a stimulus for progressive work in the basic sciences, such as transplantation immunology and its application to in vivo models. Operative techniques and indications for small bowel transplantation are discussed in detail, and suggestions for clinical application are given.

Central Action of ACTH and Related Peptides

Editors: D. de Wied, University of Utrecht;
W. Ferrari, University of Modena

1986. VIII, 240 pages. (FIDIA Research Series).
Hard cover DM 175,–. ISBN 3-540-96367-7
Distribution rights for Italy: Liviana Press, Padova

A. Labhart, University of Zürich

Clinical Endocrinology

Theory and Practice

2nd, completely revised edition. 1986. 416 figures. Approx. 1125 pages. Hard cover DM 360,–. ISBN 3-540-15462-0

"The photographs, tables and diagrams are an outstanding feature of this work. One can only state that this book is far in advance of its competitors and will readily alter the habits for those who require a standard reference text in endocrinology." *Gynecological Investigation*

Springer

T. M. Fenchel, University of Aarhus

Ecology of Protozoa

The Biology of Free-living Phagotrophic Protists

1986. 47 figures. Approx. 200 pages. (The Brock/Springer Series in Contemporary Bioscience). Hard cover DM 94,– ISBN 3-540-16960-1
Distribution rights for the USA, Canada and Mexico: Science Tech, Inc., Madison, WI

Here is a comprehensive treatment of the functional biology, ecology, and natural history of free-living protozoa. Emphasis is placed on the importance of protozoa in the food chains of natural ecosystems, and this role is discussed within the framework of the functional properties of these organisms.

J. Brachet, H. Alexandre, Free University of Brussels

Introduction to Molecular Embryology

2nd, totally revised and enlarged edition. 1986. 77 figures. Approx. 240 pages. (Heidelberg Science Library). Soft cover DM 39,–. ISBN 3-540-16968-7

Contents: From Descriptive to Molecular Embryology. – How Genes Direct the Synthesis of Specific Proteins in Living Cells. – How Eggs and Embryos are Made. – Gametogenesis and Maturation: The Formation of Eggs and Spermatozoa. – Fertilization: How the Sleeping Egg Awakes. – Egg Cleavage: A Story of Cell Division. – Molecular Embryology of Invertebrate Eggs. – Molecular Embryology of Amphibian Eggs. – Molecular Embryology of Mammals. – Biochemical Interactions Between the Nucleus and the Cytoplasm During Morphogenesis. – How Cells Differentiate.

Concepts in Viral Pathogenesis II

Editors: A. L. Notkins, Bethesda, MD; **M. B. A. Oldstone**, La Jolla, CA

1986. 31 figures, 12 tables. XX, 428 pages. Hard cover DM 110,–. ISBN 3-540-96322-7

This extension of the well-received first volume is in the same format with 61 minireviews by leading researchers on new topics in viral structure and function, oncogenes, viral immunology, human disease, and antiviral vaccines. It offers ready access to current knowledge on one of the most important and rapidly developing subjects in biomedical science today.

By the same editors:

Concepts in Viral Pathogenesis

1984. 26 figures. XVIII, 409 pages. Hard cover DM 106,– ISBN 3-540-90982-6

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo
Heidelberger Platz 3, D-1000 Berlin 33 · 175 Fifth Ave., New York, NY 10010, USA
28, Lurke Street, Bedford MK40 3HU, England · 26, rue des Carmes, F-75005 Paris
37-3, Hongo 3-chome, Bunkyo-ku, Tokyo 113, Japan

J. W. Moore, Vegreville, Alta., Canada

The Changing Environment

1986. 40 figures. X, 239 pages. (Springer Series on Environmental Management). Hard cover DM 86,– ISBN 3-540-96314-6

Contents: Introduction. – Environmental Issues in Less Developed Countries. – Environmental Management in Developed Countries. – Hazardous Waste. – Ground Water Contamination. – Toxic Substances in Water. – Toxic Particulates in Air. – New Fossil Fuel Technologies. – Nuclear Energy. – Acid Deposition. – Index.

Smiths' Sea Fishes

Editors: M. M. Smith, Grahamstown; **P. C. Heemstra**, Washington, D.C.

6th edition. 1986. Approx. 1800 figures, 144 colour plates. XX, 1047 pages. Hard cover DM 198,–. ISBN 3-540-16851-6 (Originally published by the Trustees of the Sea Fishes of Southern Africa Book Fund)

Describes all important aspects of the fishes of the South Atlantic and the Indian Oceans. Its wealth of colour and black-and-white illustrations facilitate the quick identification of a species and convey the beauty of the living specimen.

Ecophysiology of Spiders

Editor: W. Nentwig, University of Regensburg

1986. Approx. 133 figures. Approx. 465 pages. Hard cover DM 278,–. ISBN 3-540-17034-0

Ecophysiology encompasses the wide range of physiological characteristics that determine the ecology of a species, thus providing a suitable framework for this first book on spider physiology. Experts have written short reviews and original research reports on such diverse topics as thermal relationships, gland physiology, feeding ecology, parasites, and biorhythms.

Prices are subject to change without notice

3570/4/1

Springer

Reader Service No.59

Blackwell Scientific Publications

A Genetic Switch: Gene Control and Phage λ

M. **Ptashne**, Professor of Biochemistry and Molecular Biology, Harvard University

A Genetic Switch has been written by one of the most eminent names in the field of molecular genetics as an introduction to gene function - specifically the genetic repressor. Understanding the repressor function is essential to an understanding of biochemical activity within the gene as it is fundamental to the process by which growth and development are regulated at the molecular level. The book is intended as a basic introduction to repressor function for both advanced and undergraduate and graduate level courses in genetics and biochemistry. It will also be essential reading for researchers and students of phage, viral and general genetics.

1986. 138 pages, 20 illustrations. Paper, £14.95
0 86542 315 6 *Blackwell*

Essential Genetics

P. J. Russell BSc, PhD, Associate Professor of Biology, Reed College, Portland, Oregon

Early 1987. 326 pages, 277 illustrations. Paper, about £13.50
0 632 01602 7 *Blackwell*

In this book, based on an earlier book by the author in the *Lecture Notes* series, this exciting and fast moving field is condensed into a brief but balanced review. Principles of genetics are presented in an up to date and readable way, together with an explanation of both classical and modern genetics experiments which enable the reader to gain an understanding of the methodology of genetics. The format is logical and concise, with chapters on the molecular aspects of genetics, transmission genetics, regulation of gene expression, and extranuclear inheritance. Topics of particular interest included are a discussion of rapid DNA sequencing and gene cloning, regulation of gene expression, the action of steroid hormones and a discussion of the genetics of populations. Each chapter contains exercises for the student, with answers supplied at the end, and there is a useful glossary for quick reference.

Blackwell Scientific Publications Ltd
Osney Mead, Oxford OX2 0EL
Telephone (0865) 240201

Reader Service No.23

To Appear in 1987

An important new forum for researchers in this multidisciplinary field . . .

Journal of Paleoclimatology

Editors
Alan D. Hecht, *NOAA, 11400 Rockville Pike, Rockville, MD20852, USA*
Tom M. L. Wigley, *Climatic Research Unit, University of East Anglia, Norwich, NR4 7TJ, Great Britain*

Scope

The *Journal of Paleoclimatology* (JPC) provides a focus for all aspects of research in paleoclimatology, a common meeting place for the otherwise widely dispersed disciplines that contribute to this multidisciplinary field. Original work on virtually any aspect of paleoclimatology falls within the scope of the journal. New paleoclimatic evidence from all forms of proxy data on all time scales will be welcome. JPC aims to encourage new techniques for the quantification of proxy climatic data, new methods of analysis and interpretation, and the use of numerical climate models, particularly where this leads to a better understanding of the causes of climatic change.

Subscription Information
ISSN: 0920–3044
1987, Volume 1 (4 issues)
Institutional rate: Dfl. 226.00 / US$ 94.00
Private rate: Dfl. 110.00 / US$ 46.00
Prices include postage and handling

D. Reidel Publishing Company
A Member of the Kluwer Academic Publishers Group
P.O. Box 17, 3300 AA Dordrecht, The Netherlands
101 Philip Drive, Norwell, MA 02061, U.S.A.
Falcon House, Queen Square, Lancaster LA1 1RN, U.K.

Reader Service No.62

Feathered flights of fancy

Tom Kemp

Archaeopteryx, The Primordial Bird: A Case of Fossil Forgery. By Fred Hoyle and Chandra Wickramasinghe. *Christopher Davies, PO Box 403, Sketty, Swansea SA2 9BE, UK:1986. Pp.135. £10.95.*

THE eminent British astronomers Fred Hoyle and Chandra Wickramasinghe have greatly exercised the vertebrate palaeontology community with their assertion that the famous reptile/bird fossil, *Archaeopteryx*, is a fake. If true, the claim would match the enormity of the fraud associated with Piltdown man. Never having studied any of the specimens of *Archaeopteryx*, nor staked my reputation upon its genuineness, I can cheerfully accept the possibility in principle. Briefly, what is supposed to have happened is that the fossil of a small dinosaur was taken, shallow areas of matrix gouged out and replaced by a paste made from ground-up matrix, and then modern feathers pressed onto the paste to create the imprints now seen.

The authors go about the task of proving their theory with all the comprehensive enthusiasm of fictional detectives, although of the Inspector Clouseau rather than Sherlock Holmes variety. For, whether through perversity or stupidity I do not know, they exhibit a staggering ignorance about the nature of fossils and fossilization processes. Others have commented in detail on the exact misunderstandings (see the paper by A. J. Charig *et al.* in *Science* **232**, 622–626;1986). In general terms, Hoyle and Wickramasinghe draw our attention to such points as changes in the texture and coloration of different parts of the matrix, failure of the part and counterpart slabs to match precisely, the "impossibly" large areas of feather impressions exposed on the slab and the "double impression" of some of the feathers. These are illustrated by the excellent photographs, although not always as convincingly as the text claims. The point is, however, that if features such as these are taken as evidence for forgery, then I do not believe that I have ever seen a genuine vertebrate fossil in my life! They all reflect quite common, perfectly explicable aspects of the process of formation of a fossil.

Two things interest me about this otherwise worthless book: why it was written, and what effect it might have. The motive seems to relate to the authors' own theory of evolution, described very briefly in the first chapter. It seems that major evolutionary change occurs as a result of "genetic storms", whence showers of virus particles from outer space invade the Earth and become incorporated into the genomes of living organisms. That this theory blithely ignores the multitude of anomalies it creates out of phenomena at present perfectly well explained by conventional theory is set aside, despite the glib invocation of Occam's razor. Apparently, periods of mass extinction as well as major evolutionary change result from large genetic storms, and therefore the origin of a group such as the birds must have coincided with a mass extinction. The nearest relevant-looking mass extinction was at the end of the Cretaceous; therefore there could not have been any birds in the Upper Jurassic; therefore *Archaeopteryx* must be a fake (and the toothed birds of the Cretaceous misidentified to boot). To support this conclusion the most outrageously improbable conspiracy theory has to be invoked. For more than a century palaeontologists seem to have been involved in a Watergate style cover-up; heavy hints that the present staff of the British Museum (Natural History) are still cheating and lying are not excluded. Do astronomers really apply this standard of reasoning in their own field?

As to the effect of the book, I am afraid that it has a superficial appeal and will undoubtedly gain currency amongst Creationists ("Now that we know *Archaeopteryx* to be a fake..."). Thus will science, the scientific method and the cultural importance of trying to understand our world suffer yet another blow. At best this book is mischievous; at worst it might well be described as evil, because it betrays that morality of objectivity that science cherishes. Perhaps *Archaeopteryx* is a fake. Certainly the possibility should be investigated. But it should be done by those who actually understand fossils, fossilization and fossil preparation, not by a couple of people who exhibit nothing more than a Gargantuan conceit that they are clever enough to solve other people's problems for them, when they do not even begin to recognize the nature and complexity of those problems. ☐

Tom Kemp is Curator of the University Museum, Parks Road, Oxford OX1 3PW, UK, and a Lecturer in the Department of Zoology, University of Oxford.

New in paperback

- *Biophilia: The Human Bond with Other Species* by Edward O. Wilson (reviewed in *Nature* **312**, 205; 1984). Publisher is Harvard University Press, price is $7.95, £5.95.
- *Hawks, Doves, and Owls: An Agenda for Avoiding Nuclear War* edited by Graham T. Allison, Albert Carnesale and Joseph S. Nye, Jr (reviewed in *Nature* **318**, 415; 1985). Publisher is W.W. Norton, price is £5.75, $6.95.
- *Quantum Field Theory* by L. Ryder (reviewed in *Nature* **315**, 610; 1985). Publisher is Cambridge University Press, price is £15, $24.95.

Ornithology from Princeton

A Guide to the Birds of Colombia
Steven L. Hilty and
William L. Brown
Plates by *Guy Tudor*
H. Wayne Trimm, John Gwynne, Larry McQueen, John Yrizarry, and *P. Prall*
Line Drawings by *Michel Kleinbaum* and *John Gwynne*

Describing all of Colombia's birds, Steven Hilty and William Brown bring together information on one of the world's largest avifaunas — nearly 1,700 species. Over half of all the species of birds in South America are included, thus making the book useful in regions adjacent to Colombia, as well as in the country itself.

836 pages. 56 color plates, 13 halftone plates, 99 line illus., 2 maps. Paperback edition sewn for extra strength.
P: $42.50. C: $95.00

Birds of New Guinea
Bruce M. Beehler, Thane K. Pratt, and *Dale A. Zimmerman*
Illustrations by *Dale A. Zimmerman* and *James Coe*

This book is the first definitive ornithological field guide to New Guinea, the world's second largest island and one of the richest tropical environments in the Old World. Because of its isolation, New Guinea's biota includes spectacular radiations of unusual plants and animals (of which the birds of paradise are perhaps best known). This book treats all of the more than 700 species of birds recorded from the region, illustrating more than 600 of them in forty-seven fine color plates and eight black-and-white halftone plates.

293 pages. 47 color plates, 8 halftone plates, 16 line illus., 4 maps plus endpapers. Paperback edition sewn for extra strength.
P: $37.50. C: 65.00

At your bookstore or
Princeton University Press 41 William Street Princeton, NJ 08540

Reader Service No.2

Putting flesh onto the bones

Malcolm C. McKenna

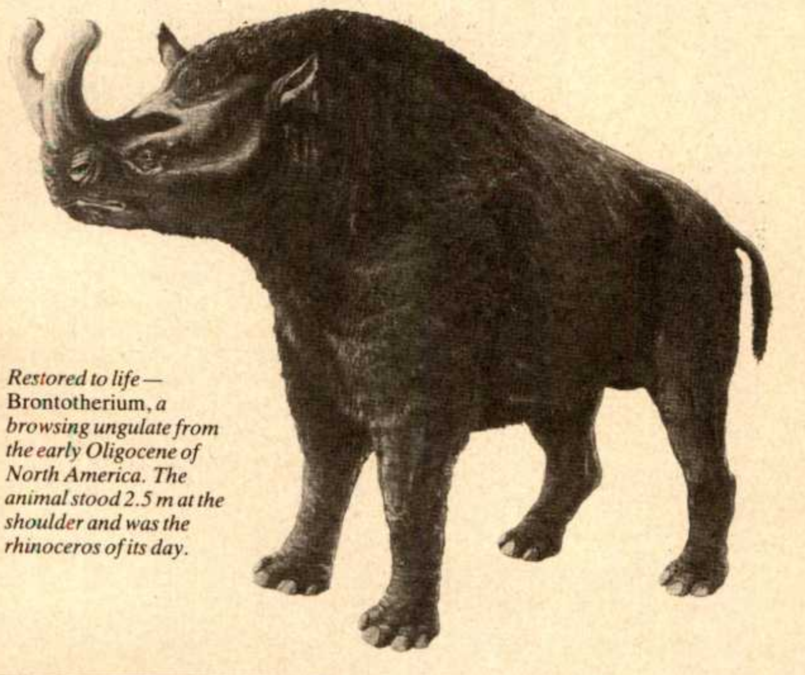
Mammal Evolution: An Illustrated Guide. By R. J. G. Savage and M. R. Long. *British Museum (Natural History)/Facts on File:1986. Pp.264. £17.50, $35.*

For most fossil mammals, dry bones have been brought to life inadequately by professional palaeontologists. Among extinct animals, the dinosaurs have stolen most of the limelight, and much of the rich store of research-level information about mammalian evolution reaches the public only as background material to the latest discoveries of rival anthropologists. Although there is a great story to tell about mammalian evolution, and a great opportunity to discuss how the scientific method can be applied to the subject, the good books on mammalian evolution over geological time have long been out of print or are badly outdated. Palaeontologists are bombarded by requests about the appearance in life of extinct mammals, and how they changed with time, but we possess few attempted reconstructions to show to the public; those that we do have are generally clones of previous attempts. Although professionals sometimes wince at reconstructions of fur colour patterns and other indefensible extrapolations from what is actually preserved in the fossil record, such images are undeniably useful for non-specialists. Life is breathed into the subject.

In this book, Michael Long (artist) and R. J. G. Savage (palaeontologist) do bring the evolution of mammals to life, from their poorly known beginnings in the Mesozoic to the richly known record of the Cenozoic. The book combines evocative restorations with spritely text. Many of the illustrations are new and plausible reconstructions of unfamiliar beasts; others clearly owe a debt to previous renderings. Some are outrageously funny, such as the human-faced *Thylacosmilus nicki*, or the contrite *Gigantopithecus* and wistful *Lystrosaurus*. The text is well written and pitched at the right level to maintain interest, mixing solid information with interesting references and speculation. It starts with a quotation from Shelley's *Prometheus Unbound* and ends with Dale Russell's silly illustration of a hypothetical ET-like "dinosauroid hominid". Between these there is much food for thought for non-specialists. There are a number of minor errors, but these can be weeded out for the future editions that will surely follow.

Our understanding of mammalian evolution is itself evolving, as more complete specimens are found and, perhaps more importantly, as methods of analysis improve. Although I found the book rewarding, one criticism is that the authors have missed the chance to move up from Kiplingesque scenarios to epistemologically more mature ground. After all, the British Museum (Natural History) is leading the way in injecting scientific method into its exhibits about evolution. To complement that effort, the book could have moved from exciting animals to greater emphasis on exciting problems. ☐

Malcolm C. McKenna is Frick Curator in the Department of Vertebrate Paleontology, American Museum of Natural History, New York, New York 10024, USA.



From *Mammal Evolution*

Restored to life— Brontotherium, *a browsing ungulate from the early Oligocene of North America. The animal stood 2.5 m at the shoulder and was the rhinoceros of its day.*

Edges of knowledge

A. Hallam

Controversy in Victorian Geology: The Cambrian–Silurian Dispute. By James A. Secord. *Princeton University Press:1986. Pp.363. $49.50, £33.10.*

Though geologists for the most part do not exhibit much interest in the history of their subject, there are few, in Britain at least, who do not know a little about the battle of those nineteenth-century giants Adam Sedgwick and Roderick Murchison. What had begun in friendly collaboration gradually developed over a quarter of a century into a bitter feud resulting in mutual estrangement. Like many stories known only in outline, however, the Cambrian–Silurian controversy is surrounded by misconceptions or is not properly appreciated for its most essential characteristic: it was a debate about the proper way to conduct a particular type of science, rather than a mere territorial or priority dispute. With an abundance of letters, notebooks and accounts of meetings available for study, the subject has long been ripe for the sort of book that Secord has written, a detailed, scholarly analysis using modern insights into the sociology of Victorian science.

Through the application of the novel technique of correlating strata by fossils, by the early 1830s great strides had been made in establishing both a relative timescale and producing a reliable geological map for the structurally uncomplicated younger rocks in Britain. The ancient "grauwacke" or transition rocks of Wales, however, were a geological *terra incognita* that provided a kind of new frontier challenge. Sedgwick started work in North Wales and Murchison in the Welsh Borders, which led respectively to their proposing the Cambrian and Silurian systems, with a mutually agreed boundary being established. After some years it became apparent that Sedgwick's Upper Cambrian corresponded with Murchison's Lower Silurian. The newly created Geological Survey adopted Murchison's classification and Sedgwick became increasingly isolated. But the turning point came with the recognition of a major unconformity between Murchison's Lower and Upper Silurian, and the discovery of "primordial" fossils at a stratigraphic level well below that in which anything had previously been found. In the 1850s and 1860s many leading geologists switched their allegiance to a three-fold classification of what we now call the Lower Palaeozoic, and the dispute effectively came to an end in 1879 when Charles Lapworth proposed the term Ordovician for Murchison's Lower Silurian.

Secord argues that "Cambria" and

"Siluria" were conceived of and valued as labels for scientific properties and signals of accomplishment in the quest for originality, with Sedgwick fighting for the Cambrian as a compensation for his failure to complete a major geological book and the more urbane but autocratic Murchison exhibiting a kind of imperialist expansionism in spreading the Silurian both across the world and down to the base of the oldest recognizable fossil-bearing rocks. Whereas Murchison laid emphasis on fossils as the only sound basis for correlation, Sedgwick thought it imperative to establish rock sections with a variety of physical criteria in a way that exemplified the principles of induction that he taught his students at Cambridge; the vehement defence of his Cambrian territory was a kind of moral crusade for the principles of truth and justice as he saw them.

Throughout the dispute, we perceive a characteristic Victorian preoccupation with facts and their classification, rather than with the testing of theories involving process and time. Secord brings out clearly the influence of various social factors in affecting the position which particular people adopted, such as personal competition, the character of the institution to which they belonged, the opportunities for professional advancement and so on. Furthermore, the nature of the underlying assumptions, and indeed the whole conceptual framework, had an influence on what were recognized as "objective" facts.

Secord adopts a straightforward narrative approach and has produced a well-written, sometimes absorbing account which demonstrates convincingly the heuristic value of controversy in promoting scientific advance. What I missed, however, was more of the authentic voice of the leading disputants. Sedgwick was by all accounts a lively and engaging, if somewhat choleric, person who expressed himself with great force and eloquence. It would have been gratifying to see more evidence of this through the inclusion of extended extracts from his correspondence and notebooks, and to contrast it with the more restrained but dogmatic tone of Murchison who is likewise poorly served by quotation. □

A. Hallam is Lapworth Professor of Geology, University of Birmingham, PO Box 363, Birmingham B15 2ET, UK.

Top people read...

John C. Marshall

The Harvard Guide to Influential Books: 113 Eminent Harvard Professors Discuss The Books That Have Shaped Their Thinking. Edited by C. Maury Devine, Claudia M. Dissel and Kim D. Parrish. *Harper & Row:1986. Pp.300. Hbk $18.95; pbk $7.95.*

WHAT do *Rats, Lice, and History, Dynamic Administration, Amateur Telescope Making* and the *Adventures of Pinocchio* have in common? The answer is that each was chosen by *one* of "113 eminent Harvard professors", who here "discuss the books that have shaped their thinking". I emphasize "one" because the most striking characteristic of *The Harvard Guide to Influential Books* is how few choices the contributors hold in common. Between them, the distinguished professors refer to some 560 works, most of which I (and you) have not read. I accordingly thought that the best way to educate myself would be to pick out the *really* influential books, say those that at least five contributors mention. We are now left with *The King James Bible* (seven listings), *War and Peace* (six) and *The Republic, The Federalist Papers* and *The Adventures of Huckleberry Finn* (five each). If authors with listings for multiple books are counted, Winston Churchill, Sigmund Freud, Thomas Mann, William Shakespeare and Max Weber just manage to creep in.

To catch a scientist we must drop to four citations for *On the Origin of Species*. Physicists and chemists are conspicuous mainly by their absence. No one seems to have been much impressed by Galileo, Newton or Max Planck, nor inspired by Lavoisier, Dalton or Mendeleev. Gone are the days when eminent Harvard mathematician Tom Lehrer concluded his song of the elements with the immortal lines: "These are the only ones of which the news has come to Harvard. There may be many others but they haven't been discarvard". Einstein is cited once, by Nicolaas Bloembergen, for his 1938 book (with Infeld) *Physik als Abenteuer der Erkenntnis*. Other physicists enter either as philosophers of science or as (proto-) molecular biologists. (No prizes for guessing Schrödinger's *What is Life?* in the latter category.) Mathematicians figure only as statisticians and "decision theorists" (*vide* Von Neumann and Morgenstern's *Theory of Games and Economic Behavior*). No one's life was made more (or less) uncertain by Heisenberg's principle; no one was converted to existentialism by Gödel's theorem.

In part, the obvious omissions result from the editors' admonition that contributors should not merely recite "Western Civilization's Great Books", but rather tell of the printed matter that *actually* influenced them. Much of the compilation's charm arises from this constraint. Thus architect Moshe Safdie engagingly praises Rudofsky's *Architecture without Architects;* E.O. Wilson confesses that, at age sixteen, Lysenko's *Heredity and Its Variability* enchanted him with its "radical and effective challenge to conventional science"; and, best of all, Mary Chatfield cites "any good cookbook" as evidence that "the rational is still achievable in the most chaotic of worlds". Occasionally one is perplexed by the reasons given for a particular choice. Historian Bernard Bailyn admires Ernest Jones's *Life and Work of Sigmund Freud* as "a morality tale of... heroic achievements", but spoils the effect by concluding "And it happens to be true". For psychologist Richard Herrenstein, *The Trial* and *The Castle* "captured the lunacy and futility of individuals struggling with bureaucracies". Poor Frank: when will people learn that Kafka is not kafkaesque?

None of the faculty has been shaped by the *Talmud*, the *Life of the Buddha* or the *Koran;* indeed, insofar as the "East" (either near or far) exists at all for the contributors it is usually in the form of economic competition with Japan (*vide*, for example, *Japan as Number One: Lessons for America*). Piety comes in predominantly paternal form, although Diana Eck (a professor of comparative religion) writes thoughtfully about Mary Daly's *Beyond God the Father: Toward a Philosophy of Women's Liberation*, and Sidney Verba (director of the University Library) recommends Chaim Grade's memoirs of the *shtetl* in *Rabbis and Wives*. Philosophy tends to the conservative; when Aristotle, Hayek, Hegel, Mill and Popper turn up it is more for their politics than their metaphysics. The spirit of the Law School, the Business School and the Kennedy School of Government dominates the enterprise. *Das Kapital* receives a derisory two votes, *The General Theory of Employment, Interest, and Money* scores zero. Proudhon, Bakunin and Prince Kropotkin fail to make the grade, whilst *Explaining the Cuban Missile Crisis, America's Competitive Edge* and Machiavelli's *Prince* get straight A's. Balance is only marginally restored by educationalist Francis Keppel, who cites *The New York Times*, a choice that serves to remind us that Britain has not had a decent daily since the *Manchester Guardian* dropped the *Manchester*.

Reviewers often decline to take compilations such as this seriously. *Au contraire*, I found the *Harvard Guide* fascinating and deeply informative. My only regret is that space was not found for the book-lists of eminent Harvard secretaries — the people who really keep an institution moving. I eagerly anticipate a series of such volumes from major universities. They would provide a far better guide to advising one's children on which centres of higher learning to apply to than do the usual bland prospectuses. □

John C. Marshall is in the Neuropsychology Unit, part of the Neuroscience Group at the Radcliffe Infirmary, Oxford OX2 6HE, UK.

ORSTOM

INSTITUT FRANÇAIS DE RECHERCHE SCIENTIFIQUE
POUR LE DÉVELOPPEMENT EN COOPÉRATION

Guide des étoiles de mer, oursins et autres échinodermes du lagon de Nouvelle-Calédonie. Handbook of the sea-stars, sea-urchins and related echinoderms of New-Caledonia lagoon. A. GUILLE, P. LABOUTE, J.-L. MENOU. Partly bilingual : French, English. 1985. 240 p., 224 photos coul. 350 F.

Cloffa 1. Check-list of the freshwater fishes of Africa. 1984. 410 p. Musée royal de l'Afrique centrale/ORSTOM. 130 F.

Cloffa 2. Check-list of the freshwater fishes of Africa. 1986. 520 p. Musée royal de l'Afrique centrale/ORSTOM. 140 F.

Mangroves du Sénégal et de la Gambie. Ecologie - Pédologie - Géochimie - Mise en valeur et aménagement. C. MARIUS. 1985. 368 p., 10 pl. photogr. 135 F.

Changements globaux en Afrique durant le Quaternaire. Passé - Présent - Futur. INQUA - ASEQUA. Symposium international. Dakar 21-28 avril 1986. 1986. 528 p., some articles in English. 300 F.

Monographie hydrologique du fleuve Niger. Ministère des Relations Extérieures. Secrétariat d'Etat à la Coopération. ORSTOM. **Tome I - Niger supérieur.** 1986. 396 p., 4 cart. ann. noir, 7 microfiches. **Tome II - Cuvette lacustre et Niger moyen.** 1986. 506 p., 6 cart. ann. noir. (the 2 volumes : 350 F).

Pharmacopées traditionnelles en Guyane. Créoles, Palikur, Wayãpi. P. GRENAND, C. MORETTI, H. JACQUEMIN. 496 p., 156 photogr. coul., 49 traits. *(forthcoming)*

EDITIONS DE L'ORSTOM
70 - 74 Route d'Aulnay
93140 BONDY

Reader Service No.17

Daniel by himself

Jacquetta Hawkes

Some Small Harvest: The Memoirs of Glyn Daniel. *Thames & Hudson:1986. Pp.448. £12.95.*

GLYN Daniel, with characteristic sangfroid, has chosen the tile of *Some Small Harvest* for one of the bulkiest autobiographies of recent times: 448 pages and over 50 illustrations. The first photograph shows a captivating little boy with bright, enquiring eyes, perhaps capable of impudence, and stoutly booted for the successful passage through worldly life that is to be chronicled in such loving detail.

Here I must declare an interest. Glyn Daniel went up to St John's College, Cambridge, during that dawn of modern archaeology, the early 1930s, in the very year that I had graduated. He came from a small village in South Wales, I from a home within half a mile of St John's. We both read for the new archaeological tripos and were to pursue the subject, he with infinitely greater diligence, for the rest of our lives. So it is that I share his memories of very many of the individuals and the events great and small which fill his book. It is hard for me to judge whether others, who must read without the glow of nostalgia, will find it as appealing as I do.

Glyn Daniel not only has an extraordinarily retentive memory but through 40 years has kept a journal which arms him with a record that is "contemporary and explicit". What is striking about the resulting life story is its range and variety. It is an exceptional man who, emerging from an unworldly yeoman background, could become a Wing Commander in India (as head of Central Photographic Interpretation), a brilliantly successful editor of *Antiquity* and of a long series of books of the best *haute vulgarisation*, an adventurous Steward of St John's, a writer on food and drink and of detective fiction — and then, when many felt that all this might have weakened his scholarly repute, carry off the much-honoured Disney Professorship of Archaeology at Cambridge.

Daniel writes of these achievements, interspersed with amusing episodes and anecdotes, as though he took life easy as the grass grows on the weirs, whereas in reality there was within him an iron determination to succeed and a rare capacity for hard work. He has such an enthusiastic interest in his own doings that he perhaps sometimes falls into triviality and self-indulgence; there are, for instance, too many recipes and wine lists, and accounts of lectures and broadcasts — yet for much of the way he carries one along on a tide of his own enjoyment.

For readers of *Nature*, most attention should go to Daniel's archaeological researches and judgements over half a century. Here there is some disappointment. He has published several histories of the subject and can be excused from all such generalities. Yet it would, I think, have given more unity to the autobiography if he could have interwoven the development of archaeological knowledge, method and thought with his personal story; in particular, if he could have gone more fully into the new radiocarbon dating which appears to have nullified much of his extensive research into megalithic monuments, and also examined the present anti-diffusionist fashion. For fashion I believe it to be: how many publications have I read of late where the author is plainly forcing his conclusions into the "local development" mould. It should be added, however, that Glyn Daniel is consistently opposed to what he chooses to call "bullshit" archaeology, the outpourings of those who are honestly deluded or who deliberately indulge in wild fantasies in order to sell their books by the million.

From *Some Small Harvest*

Glyn Daniel, aged five — "stoutly booted for the successful passage through worldly life".

Standing back to judge *Some Small Harvest* as dispassionately as I can, I am convinced that in the long run it will remain valuable as a sociological source-book, a record of Cambridge and especially college life, of travel and conditions in many lands, of the flavour of twentieth-century archaeology — and a highly personal *Who's Who* of scores of the notable characters the author has known. Viewed in this light the very trivia I have criticized will be of unique interest for a future generation, and may be more than forgiven.

Jacquetta Hawkes, Littlecote, Leysbourne, Chipping Camden, Gloucestershire, GL55 6HL, UK, is an archaeologist and author of, among other books, Mortimer Wheeler: Adventurer in Archaeology *(Weidenfeld & Nicholson, 1982).*

NEW BOOKS FROM TAYLOR & FRANCIS

METABOLISM AND ACTION OF ANTI-CANCER DRUGS

Edited by: **G Powis**, and **R A Prough**

This book provides a survey, by a number of experts, of the metabolic processes of the major classes of currently known anti-cancer drugs.

0 85066 369 5 Hard 300pp
Published: 1987 £30/$60

CANCER CAUSES AND PREVENTION

Swedish Cancer Committee

This book is the result of a report commissioned by the Swedish Government in 1979, to examine the incidence of all forms of cancer and its correlation with possible causative factors. It comments on aspects of cancer diagnosis and treatment and on possible social, political, legal, and environmental measures that could reduce the incidence.

0 85066 344 5 Hard 500pp
Published: 1987 £45/$90

PROGRESS IN DRUG METABOLISM

Volume 10

Edited by: **J W Bridges, L F Chasseaud,** and **G G Gibson**

This is an established series containing up-to-date critical reviews of topics concerning the biological fate of drugs and other xenobiotics.

0 85066 372 5 Hard 300pp
Published: 1987 £35/$77

PESTICIDE EFFECTS ON SOIL MICROFLORA

Edited by: **L Somerville, M P Greaves** and **K Domsch**

Pesticides aimed at specific pests on crops end up to a large extent on the ground and, being toxic, can have significant effects on the microflora. This book contains the papers submitted to a workshop on this subject held in 1985.

0 85066 365 2 Hard 250pp
Published: 1986 £30/$60

CHEMICALS FOR ANIMAL HEALTH CONTROL

George C Brander

The aim of this book is to discuss the impact of new methods of disease control in animals and to consider many of the new fields of therapy which have become important to the farmer.

0 85066 311 3 Soft £12.50/$25
0 85066 358 X Hard £25/$50
Published: 1986 220pp

CONTROL OF MAMMAL PESTS

Edited by: **C G J Richards** and **T Y Ku**

This book deals with control, mainly with poisons, of mammal pests, especially rodents, and the methods and strategies for countering these pests.

0 85066 371 7 Hard 250pp
Published: 1986 £39/$80

CLINICAL PHARMACOLOGY OF ANTI-INFLAMMATORY AGENTS

P M Brooks, W F Kean and **W W Buchanan**

This volume provides a concise review of the process of inflammation, giving an integrated understanding of the pharmacology of analgesics, non-steroidal anti-inflammatory drugs, the non-specific anti-rheumatic drugs and the clinical treatment of inflammatory rheumatic diseases.

0 85066 331 8 Hard 160pp
Published: 1986 £16/$45

DRUGS AND DRIVING

Edited by: **J J de Gier** and **J F O'Hanlon**

The effects of drugs on drivers' performance are coming under increasing scrutiny, as they could be the cause of road accidents. Pressures are mounting for tests to be introduced in the pre-marketing examination of new medicinal drugs. This book is the result of a seminar held in 1984 to look into all aspects of this important subject.

0 85066 290 7 Hard 402pp
Published: 1986 £40/$58

HUMAN GROWTH: A MULTIDISCIPLINARY REVIEW

Edited by: **A Demirjian**

This book covers aspects of human growth and development throughout all stages of life. It comprises the proceedings of the 4th International Congress of Auxology held in 1985.

0 85066 354 7 Hard 300pp
Published: 1986 £32/$58

MECHANISMS OF DRUG ACTION

C J Coulson

A wide ranging textbook surveying the biochemical modes of action of pharmaceuticals in the body. The book deals with the main drugs currently in use, and also ones which are still at the clinical trial stage.

0 85066 376 8 Hard 200pp
Published: 1987 £14/$28

HUMAN ASPECTS OF COMPUTER-AIDED DESIGN

A Majchrzak, W Barfield, T-C Chang, R Eberts, G Salvendy.

This book discusses the technical, human factors and management aspects of CAD. It consolidates information for the designers and users of CAD systems in such a manner that its effective use should enable better design and use of CAD systems.

0 85066 364 4 Hard 280pp
Published: 1986 £25/$50

OCCUPATIONAL HEALTH AND SAFETY IN AUTOMATION AND ROBOTICS

Edited by: **K Noro**

This book contains the papers presented at a Symposium held in Japan in 1985, where researchers and practitioners discussed theories and their experiences in applying human factors to the design, planning, control and operation of human/robotic systems and industrial automation.

0 85066 351 2 Hard 400pp
Published: 1986 £35/$72

Rankine Road, Basingstoke, Hants, RG24 0PR, UK
242 Cherry Street, Philadelphia, PA 19106-1906, USA

Reader Service No.45

Mobilizing against AIDS
The Unfinished Story of a Virus
THE INSTITUTE OF MEDICINE/NATIONAL ACADEMY OF SCIENCES

'One of the most well-written, balanced, and accurate summaries of information relating to the AIDS epidemic yet written. It presents a rather sophisticated description of the immunology and virology of the disease in a very readable way.' Merle A. Sande, M.D., University of California at San Francisco and San Francisco General Hospital
£12.75 Cloth 216pp 0-674-57760-4
£6.75 Paper 0-674-57761-2

The History of Statistics
The Measurement of Uncertainty before 1900
STEPHEN M. STIGLER

'After hundreds of years our field has a much-longed-for masterwork by a scholar who combines his skills as a historian of science and a researcher in statistical inference with a gentle sense of humour . . . a gorgeous book.' Frederick Mosteller
Belknap £21.25 Cloth 410pp illus.
0-674-40340-1

The Sickled Cell
From Myths to Molecules
STUART J. EDELSTEIN

This book is the first attempt to summarize all that is known about the historical and cultural roots of sickle cell anemia and the molecular details of how it attacks humans.
£16.95 Cloth 224pp illus.
0-674-80737-5

Thunderstones and Shooting Stars
ROBERT T. DODD

This lovely and accessible book illuminates the complex science of meteoritics and examines the controversial hypothesis that meteorites have played a dramatic, albeit occasional, role in the evolution of life on Earth.
£21.25 Cloth 272pp illus.
0-674-89137-6

Research Interviewing
Context and Narrative
ELLIOT G. MISHLER

Interviews hold a prominent place among the various research methods in the behavioural and social sciences. This book presents a powerful critique of current views and techniques and proposes a new approach to interviewing.
£16.95 Cloth 208pp 0-674-76460-9

Dynamics of Marine Fish Populations
BRIAN J. ROTHSCHILD

This book provides the first coherent theory for analyzing the dynamics of fish populations and the number of fish born each year.
£31.95 Cloth 352pp illus.
0-674-21879-5

HARVARD
UNIVERSITY PRESS
126 Buckingham Palace Road London SW1W 9SD

Reader Service No.34

LES EDITIONS
INSERM

SYMPOSIA

Approaches to the therapy of sickle cell anaemia
Y. BEUZARD
S. CHARACHE
E. GALACTEROS
VOL. 141. 572 PAGES.
US$ 35. FF 196

Techniques en biorhéologie
J.-F. STOLTZ
VOL. 143. 372 PAGES
US$ 29. FF 150

De la recherche bio-médicale à la pratique des soins
CNAMTS/INSERM
VOL. 144, 806 PAGES, 150 FF

Human population genetics
E. OHAYON
A. CAMBON-THOMSEN
VOL. 142, 410 PAGES, 200 FF

SERIES
« GRANDES ENQUETES »

EDITIONS INSERM/DOIN (1)

Suicide and suicide attempts today
F. DAVIDSON
A. PHILIPPE
1986, 160 PAGES, 110 FF.

FORTHCOMING

Genetic markers in « provinces françaises »
E. OHAYON
A. CAMBON - THOMSEN

(1) Distribution DOIN Editeurs, 8, place de l'Odéon, 75006 Paris.

French/English Biomedical Thesaurus
INSERM-CDST CO-EDITION
A bilingual version of the Medical Subject Headings (NLM) and their PASCAL equivalents (CNRS)
1986, 3 VOLUMES, 2 400 PAGES
VOL. 1 : 1 000 FF
VOL. 2, VOL. 3 : 700 FF
LES 3 VOLUMES : 2 000 FF

SYMPOSIA

EDITIONS INSERM
JOHN LIBBEY EUROTEXT (2)

Cardiovascular and respiratory physiology in the fetus and neonate
M. MONSET - COUCHARD
P. KARLBERG. A. MINKOWSKI
W. OH. L. STERN
VOL. 133. 204 PAGES. 1986
£ 17. US$ 30. FF 210

Porphyrins and porphyrias
Y. NORDMANN
VOL. 134. 240 PAGES. 1986
£ 19. US$ 34. FF 250

Neo-adjuvant chemotherapy
C. JACQUILLAT. M. WEIL. D. KHAYAT.
VOL. 137. 864 PAGES. 1986
£ 48. US$ 84. FF 620

Hormones and cell regulation
(11 th european meeting)
J. DUMONT
J. NUNEZ
R.J.B. KING
VOL. 139, 304 PAGES, 1986
£ 25. US$ 44. FF 330

Control and management of parturition
C. SUREAU
J.PH. BLOT ET COLL
VOL. 151, 300 PAGES, 1986
£ 22, US$ 38, FF 290

(2) Orders from France. John Libbey Eurotext Ltd., 6, rue Blanche, 92120 Montrouge, France. Other orders, Cheque or Bank Draft in £ sterling or US $ payable to John Libbey & Company Ltd., 80/84 Bondway, London SW8 1SF, England.

INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
101, rue de Tolbiac. 75654 PARIS CEDEX 13
Tél. : 45.84.14.41

All books available at INSERM. Prepayment required - Bank cheque to the order of INSERM.

Reader Service No.35

Books from SCIENCE

The Journal of the American Association of the Advancement of Science

Aids

Papers from *Science*, 1982-1985

Edited by Ruth Kulstad

Some of the most frequently cited papers on acquired immune deficiency syndrome (AIDS) that appeared in *Science* between August 1982 and September 1985 are brought together in this unique source book on AIDS. The collection is particularly timely both because of the experimental data included and because the papers provide an excellent source for references to related work in other journals. Arranged chronologically, the 108 research papers and *Science* news reports show how far AIDS research has come, providing a brief history of the AIDS epidemic through September 1985. The research papers indicate the state-of-the-art techniques applied to studies of the AIDS virus and set out probable directions for future research. The news stories focus on some of the problems associated with investigating a disease occurring predominantly — at least in the US — with homosexual men, on the early reluctance of the government to provide adequate research funds, and on the problems that may be encountered in caring for large numbers of sick and dying patients.

An overview of AIDS research to date is given in the introduction by Dr Myron Essex, chairman of the Department of Cancer Biology, Harvard University School of Public Health.

Ruth Kulstad is a senior editor with *Science*.

1986-653 pages w/notes, tables, figures, biblio, indexes.
ISBN 0-87168-313-X (hc). £24.00.

Astronomy and Astrophysics

Edited by Morton S. Roberts

The fields of astronomy and astrophysics have exploded in recent years in the wake of technological and theoretical advances. This volume, a survey of the solar system, the structure and content of the galaxy, galaxies and cosmology , and instrumentation, is an attempt to assess the current state of knowledge of astronomy. The editor of this collection has organized 24 articles, including two Nobel Prize Lectures, originally published in *Science* between 1982 and 1984, on topics ranging from the solar system to the pulsars at the very edge of the observable universe. Each article is self-contained, yet as a whole, the volume reveals a broad, coherent, and contemporary picture of our astronomical universe. Selected for their depth of coverage and breadth of topics, these articles are of interest to the entire scientific community.

Morton S. Roberts is director of the National Radio Astronomy Observatory.

Contents (partial): Introduction. Solar System. Structure and Content of the Galaxy. Galaxies and Cosmology. Instrumentation.

1985; 383 pages w/colour plates, index.
ISBN 0-87168-311-3 (hc). £24.00

Biotechnology

The Renewable Frontier

Edited by Daniel E. Koshland, Jr

Like its 1984 predecessor, *Biotechnology and Biological Frontiers*, this new volume covers the latest and most important topics in biological research and development. The collection includes papers that appeared in *Science* after November 1983 and provides an update on state-of-the-art techniques in both fundamental and applied biotechnology.

Daniel E. Koshland, Jr, is the editor of *Science*.

July 1986; 400 pages w/tables, figures, biblio, index.
ISBN 0-87168-314-7 (hc). £24.00

Biotechnology and Biological Frontiers

Edited by Philip H. Abelson

Covering the most important topics at the forefront of biological research and development, this book addresses both fundamental research techniques and practical applications. Part 1, biotechnology, is devoted to a broad sampling of the current revolution in applied biology. Recombinant DNA and hybridomas, protein engineering, tissue culture of plants, and industrial fermentations are among the topics discussed. The emphasis in Part II, biological frontiers, is on basic research at a time when the enormous power of experimental techniques is producing answers to long-standing biological questions. Consisting of 36 outstanding papers from *Science* by 136 scientists representing industrial, university, and governmental laboratories and institutes worldwide, this volume will be of interest to researchers and students in all fields of biology, agriculture, and the health sciences.

Philip H. Abelson is the former editor and current deputy editor for engineering and applied sciences at *Science*.

1984; 550 pages w/colour plates, index.
ISBN 0-87168-308-3 (hc). £24.00

Frontiers in the Chemical Sciences

Selected and with an Introduction by William Spindel and Robert H. Simon

There has been an explosion in the field of chemistry within the last ten years resulting from powerful new research techniques and an expanding of areas of study. Many new tools and capabilities are being developed throughout the full range of analytical, physical, organic, biological, and materials chemistry. This volume details progress in these areas, including new work on ultrafast reaction dynamics, surfaces and catalysis, organometallics and conduction organics, enzyme mechanisms and other life processes, stereoselectivity and asymmetric syntheses, and catalysis.

William Spindel and Robert M. Simon are members of the National Research Council Board on Chemical Sciences and Technology.

October 1986; 400 pages w/tables, figures, notes, biblio, index.
ISBN 0-87168-312-1 (hc). £24.00

Neuroscience

Edited by Philip H. Abelson, Eleanore Butz, and Solomon H. Snyder

Discoveries made during the recent renaissance in neuroscience research in areas ranging from genetic engineering to clinical therapy are presented in this collection. Within its four sections — neuroplasticity, molecular biology, synaptic transmission, and behaviour — the volume provides an integrative treatment of brain anatomy, physiology, and chemistry and addresses fundamental questions concerning nervous system functioning. The editors have pulled together 27 seminal papers from *Science* representing the work of 84 researchers at private and federal laboratories and hospitals, integrating the disparate disciplines in neuroscience.

Philip H. Abelson is the former editor and current deputy editor for engineering and applied sciences at *Science*, Eleanore Butz is senior editor with *Science*. Solomon H. Snyder is director of the Department of Neuroscience and distinguished service professor of neuroscience, pharmacology, and psychiatry at The Johns Hopkins University School of Medicine.

Contents (partial): Introduction. Neuroplasticity. Molecular Biology. Synaptic Transmission. Behaviour.

1985; 460 pages w/colour plates, index.
ISBN 0-87168-309-1 (hc). £24.00.

INFORMATION
PUBLICATIONS
INTERNATIONAL

White Swan House
Godstone, Surrey
England.

ORDERS TO:
Wildwood Distribution Services
Unit 3, Lower Farnham Road,
Aldershot, GU12 4DY, England.

Reader Service No.49

Butterworths Biotechnology Series

Mammalian Cell Technology

William G Thilly, Department of Applied Biological Sciences, MIT, Massachusetts, USA

This new book presents an organised review of the essential elements necessary in the attempts to use fermentors instead of animals to grow animal cells and to produce physiologically active macromolecules. The underlying principle of the book is that mammalian cell fermentor productivity is the product of single cell productivity and the number of cells achieved in a fermentor. The contributors have presented an orderly overview of the important aspects of the genetics, physiology and growth of animal cells.

Hardcover 192 pages October 1986 £48

Overproduction of Microbial Metabolites

Z Vanek and Z Hostalek, both of the Czechoslovakian Academy of Sciences, Prague, Czechoslovakia

Overproduction of microbial metabolites is a complex process which can be achieved by using knowledge of control mechanisms on the genetic and metabolic levels. This volume presents an in-depth look at the various strategies used in industry to obtain preferential expression of desired metabolic pathways.

Hardcover 316 pages October 1986 £59

Further information on the Butterworths Biotechnology Series is available from the address below.

Butterworths

meeting professional needs

BOROUGH GREEN SEVENOAKS KENT TN15 8PH UK

Reader Service No.29

WHERE WILL IVF LEAD US?

In Vitro Fertilisation

past, present and future

Edited by S Fishel, *The Park Hospital, Nottingham,* and E M Symonds, *University of Nottingham*

Few topics are subject to such intense argument and so little understood.

This new book prepares you for this compelling debate. The editors have assembled the work of scientists, doctors, sociologists, philosophers, lawyers and politicians to present a comprehensive and international investigation of this fascinating field.

Contents

IVF – past and present • The infertile couple • Follicular growth, ovulation and the use of ovarian stimulants • Endocrine evaluation of the infertile woman • Ultrasound for oocyte recovery • Laparoscopy for oocyte recovery • Preparation for IVF in the laboratory • Andrology and semen preparation and preservation for IVF • Growth of the human *conceptus in vitro* • Computing in IVF • Replacement of the embryo • Pregnancies, abortion and birth after IVF • Variations in current IVF programmes – in-patient versus out-patient treatment • Caring for the IVF patient – nursing care • Caring for the IVF patient – counselling • The ethical debate • The legal situation • The current status of the Warnock Report and other international reports • The political debate • Test-tube babies – the child, the family and society • IVF – the future

IRL Press Ltd, PO Box 1, Eynsham,
Oxford OX8 1JJ, UK
IRL Press Inc, PO Box Q, McLean,
VA 22101-0850, USA

December 1986 277 pp
Soft 0 947946 50 0 £17.50/US$31.50
Hard 0 947946 95 0 £27.50/US$49.50

Reader Service No.42

ESTROGEN/ANTIESTROGEN ACTION AND BREAST CANCER THERAPY

V. Craig Jordan, editor

This volume, which brings together the leading scientific minds in the field of breast cancer research, traces the development of the understanding of estrogen action and its application to breast cancer therapy. The contributors, who include many of the principal scientists and clinicians involved in this period of research and discovery, trace the discovery of antiestrogens and their successful use as a nontoxic therapy for breast cancer. These papers will be of significant interest and value to students, scholars, and researchers in oncology, biochemistry, pharmacology, endocrinology, and cell biology.

The main goal of this book is to allow the planning of strategies for the application of research for the successful treatment of breast cancer with antiestrogen drugs.

ISBN 0-299-10480-X hb £45.00

The University of Wisconsin Press
One Gower Street, London WC1E 6HA

Reader Service No.19

NEW AUTUMN BOOKS FROM ELSEVIER

WORLD SURVEY OF CLIMATOLOGY, 1B

General Climatology, 1B

Elements of Statistical Analysis

by **O.M. Essenwanger**

With the development of electronic data processing, the field of statistical analysis has mushroomed. The availability of a vast amount of software allows the meteorologist to use statistics as a tool to a high degree and sophisticated statistical analysis of meteorological data is now possible. The author has selected a number of topics comprising a mixture of familiar and modern subjects of statistical analysis which will introduce the reader to climatological data processing.

1986 442 pages; US$ 131.00 / Dfl. 295.00;
ISBN 0-444-42426-1

Biogeochemical Processes at the Land-Sea Boundary

edited by **P. Lasserre,** *Paris, France and* **J.-M. Martin,** *Paris, France*
(Elsevier Oceanography Series, 43)

There is an intuitive belief among scientists that each estuary or lagoon is unique and different. Similarities do exist, however, and this book attempts to highlight some common properties and perhaps some innovative views on biogeochemical processes, and biological fluxes which are of central concern in the understanding of the land-sea boundary.

1986 x + 214 pages; US$ 53.25 / Dfl. 120.00;
ISBN 0-444-42675-2

Handbook of Environmental Isotope Geochemistry

edited by **P. Fritz,** *Ontario, Canada, and* **J.Ch. Fontes,** *Orsay, France*

The *Handbook of Environmental Isotope Geochemistry* comprises a series of books, each focussing on a selected environment, presenting in-depth reviews of all relevant studies and results obtained in specific research areas. The first volume, Part A, was described by one reviewer as "*everything we wanted to know about isotopes and much, much more... without doubt, a necessary reference book for any geochemist*". The second volume continues the coverage of terrestrial environment, and contains an equally impressive and authoritative collection of papers.

Volume 1: The Terrestrial Environment, A
1980 xii + 546 pages
US$ 106.75 / Dfl. 240.00
ISBN 0-444-41780-X

Volume 2: The Terrestrial Environment, B
1986 xii + 558 pages
US$ 115.50 / Dfl. 260.00
ISBN 0-444-42225-0

The Origin of Arcs

edited by **F.-C. Wezel,** *Urbino, Italy*
(Developments in Geotectonics, 21)

A collection of papers presented as distinguished guest lectures at the International Conference on "The Origin of Arcs" held at the University of Urbino on 22nd-25th September 1986, under the joint sponsorship of the European Union of Geosciences and the Italian Geological Society.

1986 xx + 568 pages; US$ 80.00 / Dfl. 180.00;
ISBN 0-444-42688-4

Terminal Eocene Events

edited by **Ch. Pomerol,** *Paris, France, and* **I. Premoli-Silva,** *Milan, Italy*
(Developments in Palaeontology and Stratigraphy, 9)

The papers in this volume provide an exhaustive inventory and description of the most complete sedimentary sequences across the Eocene-Oligocene Boundary (EOB) from all over the world, and present a synthesis of the biotic and chemico-physical events detected at the Eocene-Oligocene transition.

1986 xii + 414 pages; US$ 69.00 / Dfl. 155.00;
ISBN 0-444-42623-X

P. O. BOX 211 • 1000 AE AMSTERDAM • THE NETHERLANDS
ELSEVIER
O. BOX 1663 • GRAND CENTRAL STATION • NEW YORK • NY 1016

Reader Service No.60

The Life of Science from Chicago

SEWALL WRIGHT AND EVOLUTIONARY BIOLOGY

WILLIAM B. PROVINE

"An excellent piece of work. The emphasis is, as it ought to be, on Wright's contribution to science." —Ernst Mayr, Harvard University

Cloth $30.00 £25.50 562 pages
Science and Its Conceptual Foundations series

EVOLUTION

Selected Papers

SEWALL WRIGHT

Edited and with Introductory Materials by **William B. Provine**

All of Sewall Wright's published papers on evolution up to 1950, and a few published later, are gathered in this volume.

Paper $25.00 £21.25 654 pages
Library cloth edition $70.00 £59.50

INTELLECTUAL MASTERY OF NATURE

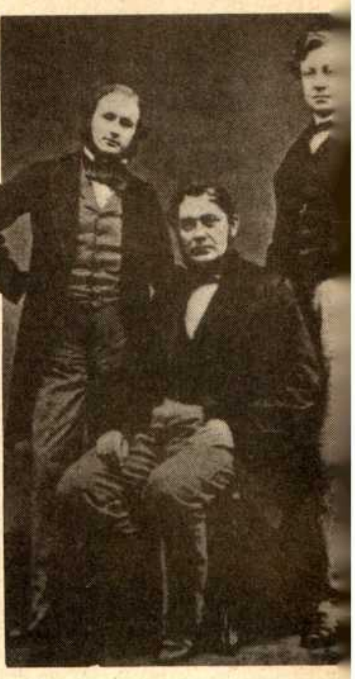
Theoretical Physics from Ohm to Einstein

CHRISTA JUNGNICKEL *and* RUSSELL McCORMMACH

Volume 1: *The Torch of Mathematics, 1800–1870*
Cloth $55.00 £46.75 378 pages
1 map

Volume 2: *The Now Mighty Theoretical Physics, 1870–1925*
Cloth $65.00 £55.25 456 pages
36 halftones, 1 map

"Few books in the history of science have been more eagerly awaited than Jungnickel and McCormmach's study of theoretical physics in Germany, and the result does not disappoint."—R. Steven Turner, *Science*



ELECTROMYOGRAPHY FOR EXPERIMENTALISTS

GERALD E. LOEB *and* CARL GANS

A complete and very practical hands-on guide to the theoretical and experimental requirements of electromyography, particularly in the biological sciences.

Paper $22.00 £18.75 394 pages
40 halftones, 100 line drawings
Library cloth edition $60.00 £50.95

THE SHAKY GAME

Einstein, Realism, and the Quantum Theory

ARTHUR FINE

An account of Einstein's realism and his opposition to quantum theory, *The Shaky Game* also examines the realism-antirealism debate in philosophy of science and proposes an original solution based on the *natural ontological attitude.*

Cloth $25.00 £21.25 198 pages
Science and Its Conceptual Foundations series

THE BIRDS OF CANADA

Revised Edition

W. EARL GODFREY

Illustrated by John A. Crosby and S. D. MacDonald

Thoroughly revised with new data and additional illustrations, this bestselling, comprehensive reference book covers all bird species known in Canada from historic times to the present.

Cloth $39.95 £33.95 596 pag
74 color plates, 107 line drawing
398 maps
Distributed for the National Museums of Canada/National Museum of Natural Sciences

Now in paper

THE LYSENKO AFFAIR

DAVID JORAVSKY

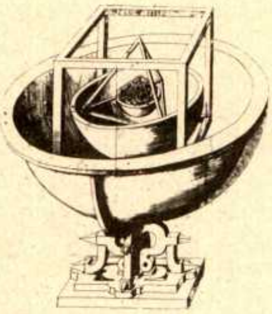
$15.95 £13.50 472 pages

IN THE RAINFOREST

Report from a Strange, Beautiful, Imperiled World

CATHERINE CAUFIELD

$9.95 320 pages 1 map



MEASURING THE UNIVERSE

Cosmic Dimensions from Aristarchus to Halley

ALBERT VAN HELDEN

$8.95 £7.50 212 pages
13 line drawings

THE ECOLOGICAL WEB

More on the Distribution and Abundance of Animals

H. G. ANDREWARTHA *and* L. C. BIRCH

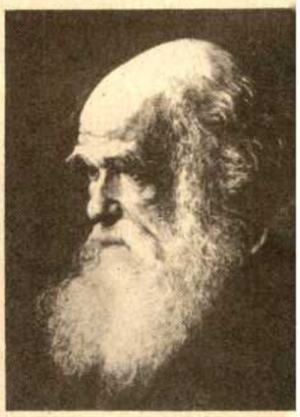
$19.95 £16.95 520 pages
12 halftones

STUDIES ON THE STRUCTURE AND DEVELOPMENT OF VERTEBRATES

EDWIN S. GOODRICH

With a new Foreword by **Keith S. Thomson**

$32.50 872 pages
754 diagrams



DIFFERENT FORMS OF FLOWERS ON PLANT OF THE SAME SPECIES

CHARLES DARWIN

With a new Foreword by **Herbert G. Baker**

$13.95 £13.95 384 pages
Illustrated throughout

The University of Chicago PRESS

Orders and inquiries for North and South America, Asia, Israel, and the Antipodes to:
5801 South Ellis Avenue
Chicago, Illinois 60637 USA

Orders and inquiries for Europe, Africa, and the Middle East to:
126 Buckingham Palace Road
London SW1W 9SD, England

Prices in US dollars and UK pounds

Reader Service No.43

Occupied territory

Randall White

The Palaeolithic Settlement of Europe. By Clive Gamble. *Cambridge University Press:1986. Pp. 469. Hbk £40, $65; pbk £15, $24.95.*

ACCORDING to current evidence, Europe was first occupied by hominids 1.5 million years after the first Palaeolithic toolmakers emerged in Africa. There is much that remains to be known about this successful expansion of our species into northern latitudes, but there is also much information presently available. Research on the Palaeolithic settlement of Europe (roughly 750,000 – 12,000 years ago) is published in so many languages, and is based upon such a variety of theories, that it is difficult for any scholar to have a clear overall view of the subject. Clive Gamble's drawing together of knowledge and references from such an intellectual mosaic is in itself a great achievement.

It must be said, however, that the framework on which the book is organized is puzzling, if not idiosyncratic. The first chapter, which begins with a history of European Palaeolithic studies, concludes with a strong advocacy of Binfordian methodological premises. While Binford has had a very positive influence on European Palaeolithic studies, it is a little surprising to find his views uncritically repeated virtually word for word.

The following chapter, "Hunter–Gatherer Regional Systems", takes a detailed look at regional and inter-regional adaptations among contemporary hunter–gatherers. At the end of it, Gamble follows Robert Foley in proposing rejection of the site as an analytical unit for the study of prehistoric hunter–gatherers. This has become a popular position recently; but it deserves much closer scrutiny in light of the fact that sites do correspond to palimpsests of past living *communities*, a concept critical to regional studies of living peoples. I found this chapter especially troublesome in its heavy emphasis on 1960s ecological thinking. The result is a form of ecological functionalism which views the social world of hunter–gatherers exclusively with reference to the imputed functions of "alliance" and "information strategies". Anthropological theory in the 1980s is not what it was in the 1960s, and for Gamble to cite Marshal Sahlins's ideas of 1968, without acknowledging Sahlins's refutation of the ecological paradigm, betrays an unduly selective approach.

Chapter 3 is a complex presentation of data on environments and resources. Gamble breaks Europe down into nine sub-regions, upon which are based the inter-regional comparisons in the rest of the book. Discussion of glacial chronology is detailed and excellent. Also here is a listing of the key anatomical and behavioural attributes of the main faunal species, which is very helpful but unfortunately omits certain attributes of importance to human beings, such as hide and antler quality and aggregation behaviour.

In the next, very short chapter, entitled "Technological and Typological Approaches", Gamble explores technological and typological evolution during the European Palaeolithic. The result is a severe revision of the traditional sequence in which the present Middle and Lower Palaeolithic are melded to form the "Early Palaeolithic", which comprises all industries before 35,000 years ago. The terms Lower and Middle Palaeolithic are inadequate, and Gamble's lumping of the periods together clears the decks for future analyses of much potentially interesting variation. It is all the more curious, then, to see him become a splitter when he turns his attention to the Upper Palaeolithic, creating two divisions (before and after 20,000 years BP). The later division is based upon the emerging importance of microliths after 20,000 years BP. Unfortunately, it ignores recent excavations in which Aurignacian (caminade scrapers) and Gravettian (microgravettes, backed bladelets) levels have yielded abundant microliths.

Chapter 5 is a useful, well-referenced discussion of the data base for research on Palaeolithic Europe, organized according to Gamble's newly developed temporal and geographical structure. Each region is summarized according to (1) chronology, (2) industrial groupings, (3) geographical variation and regional case-studies, and (4) burials and art (though the last of these is rather sketchy).

The sixth chapter, "Space and Subsistence", is the best in the book and probably the most comprehensive account of Palaeolithic settlement and subsistence patterns that has appeared to date. There are two shortcomings, however. First, because of its supposed circularity site-location research is summarily dismissed. Second, after having illustrated some brilliant work on intra-site spatial patterns, Gamble re-states his view that the site should not be our focus of analysis. My position is that the examination of individual sites must be used to build a regional picture of human landscape use.

"Demography and Style", the next chapter, is puzzling, because the section on demography deals almost exclusively with carnivore behaviour as a basis for research into the distribution of populations in space. The section on style and interaction relies heavily on information approaches to variability in material culture, exemplified by Gamble's own work on Gravettian female figurines. Again, in true 1960s fashion adaptive significance is attributed to form in portable and cave art; based upon a very narrow reading of the Australian literature, the latter is treated as territorial marking rather than as evidence for complex systems of meaning that help to define the very nature of human adaptation.

Chapter 8, "Society, Sediments and Settlement", emphasizes the interesting regional differences in stone-tool and site densities. Here, Gamble wishes to chip away at the notion that a universal set of Upper Palaeolithic traits, rather than local environmental exigencies, conditioned the archaeological record in all regions. This is indeed an important issue, but Gamble's density calculations need to be examined closely before being accepted. For example, his find-spot density for the Perigord is based on a surface area of 12,600 square kilometres, approximately twice the true figure. Moreover, historical differences in the intensity of site search need to be taken into consideration. It is a pity that Gamble fails to suggest the need for systematic controlled survey to control sample quality in assessing overall population increase across the transition between the Middle and Upper Palaeolithic.

In the concluding chapter, the virtues of viewing culture as an adaptive process are extolled, especially as applied to the emergence and diversification of culturally modern humans at about 35,000 years BP. Gamble seems not to recognize that those of us who view the Upper Palaeolithic as an evolutionary threshold are not anti-adaptive; rather, we are comfortable in switching back and forth between discussion of the evolution of culture and the evolution of cultures.

More disturbing in the conclusion, however, are the peculiarities that emerge in the form of three models for what human beings did when they reached Europe. The problem posed by cold northern climates is viewed primarily as a problem of finding a niche, which ignores clothing, fire and shelter as cultural evolutionary prerequisites. In any event, Gamble proceeds to propose that the earliest Europeans were adapted to the consumption of frozen animal carcasses resulting from natural deaths in autumn and winter. He seriously suggests that the wooden "spears" from Lehringen and Clacton were probes for finding dead animals beneath the snow. This is an unfortunate excursion onto the fringe.

While readers must work around some awkward theoretical and organizational impediments, the rich (though already somewhat dated) documentation and first-class presentation of this book will make it a basic reference in English-language Palaeolithic studies for years to come. ☐

Randall White is an Assistant Professor in the Department of Anthropology, New York University, New York, New York 10003, USA.

Principled man in unprincipled times

Martin J. Klein

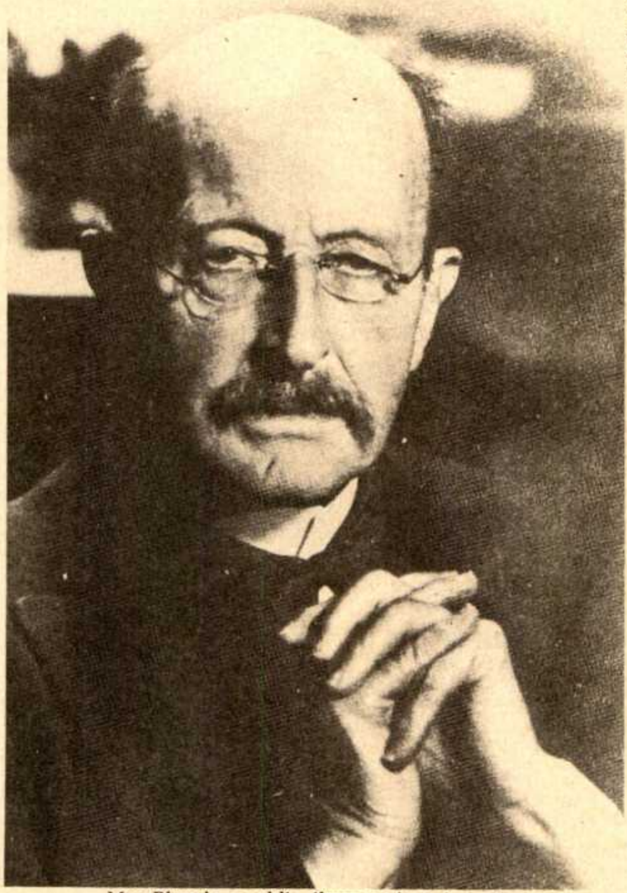
The Dilemmas of an Upright Man: Max Planck as Spokesman for German Science. By John L. Heilbron. *University of California Press: 1986. Pp.238. $16.95, £14.50.*

MAX PLANCK (1858–1947) exemplified the traditional values of the world in which he grew up, the world of a conservative, cultured German family of pastors, scholars and jurists. The very model of a classical physicist, he "embodied with rare completeness", according to Albert Einstein, the "ideal search for truth . . . a bond forever uniting scientists of all times and in all places". But his very devotion to the search for absolute scientific truth led Planck inexorably to the quantum of energy with which his name is permanently linked, the symbol of this century's nonclassical physics. His unshakable sense of responsibility led him, equally inexorably, to accept a series of positions of authority in the German academic and scientific establishments. By 1914 he had become the acknowledged spokesman for German physics, a role he could not in conscience relinquish when his country was transformed into the Weimar republic and then entered the dark times that began in 1933. As John Heilbron writes in his preface, Planck's

> struggle to come to grips with changing political circumstances, and his efforts to adjust the image of science and civic life that he developed before 1900 to the realities of twentieth-century Germany have the elements of heroic tragedy.

Heilbron's book, the first in English devoted to Planck, is an outstanding piece of work by one of our finest historians of science. Heilbron's earlier writings include an admirable study of Henry Moseley's life and work, a major book on early intellectual and institutional developments in physics, *Electricity in the 17th and 18th Centuries*, and a series of articles and monographs on a wide range of subjects in the physical science of the past four centuries. Among its other impressive features, his work always shows his mastery of "the bibliographic resourcefulness proper to the historian".

As his title implies, Heilbron's primary concern in this book is Planck's role as spokesman for science rather than the physics he actually did. Heilbron does sketch Planck's development from exclusive devotee of thermodynamics through his gradual acceptance of the need for both atomism and a statistical approach, and he shows how Planck used his hard-won theory of blackbody radiation to determine the first reliable values of the atomic constants. We see Planck, conservative in his stance towards the developing quantum theory, immediately and enthusiastically respond to the special theory of relativity and its author. Planck was one of those who brought Einstein to Berlin in 1914, and it was Planck's presence that helped to keep him there: "To be near Planck is a joy", Einstein once wrote.



Max Planck — public silence, private integrity.

Mary Evans

There is no irony in Heilbron's designation of Planck as an upright man. "Respect for laws, trust in established institutions, observance of duty, and absolute honesty" are listed as "the hallmarks of Planck's character". These were combined with genuine modesty and the clear conscience that came with his responsible adherence to his principles. Not a man for public pronouncements, Planck showed his opposition to academic anti-Semitism preceding the First World War by acting quietly to keep qualified Jewish physicists such as Emil Warburg and Lise Meitner from being denied positions on irrelevant grounds.

By 1914 Planck held a series of major positions in the German Physical Society, the Prussian Academy of Sciences, the Kaiser-Wilhelm-Gesellschaft, and, as war began, as Rector of the University of Berlin. He faced the war with mixed emotions, foreseeing "something great, along with something monstrous". Planck was one of the 93 intellectuals who signed the famous "Appeal to the Cultured Peoples of the World" in October 1914, denying German atrocities in Belgium and expressing the unity of German culture and German arms. Unlike most of his cosigners he soon moderated his position and worked to maintain the possibility of post-war international reconciliation among scientists. His 1916 open letter to H.A. Lorentz emphasized his

> firm conviction that there are domains of intellectual and moral life that lie beyond the struggles of nations, and that honourable cooperation in the cultivation of these international cultural values and, not less, personal respect for citizens of enemy states are indeed compatible with ardent love and energetic work for one's own country.

Few citizens of any belligerent nation have publicly subscribed to such enlightened views in the midst of a modern war.

Planck had to call on all his internal resources to go on with his work after his own losses in the war years: his first son was killed in action and his twin daughters both died in childbirth. Devotion to duty sustained him in this time of personal grief and national upheaval. Unlike many of his conservative colleagues he gave his loyal support to a republic whose principles were not his own. Heilbron discusses Planck's staunch support of Einstein when vicious anti-Semitic attacks were directed against both the man and his general relativity, but he stresses Planck's avoidance of public statements that he thought might encourage the politicization of science. We see Planck struggling to help restore normal international relations among scientists despite the severe restrictions imposed by the victorious Allied powers. As if these tasks and his efforts to secure adequate financial support for physicists were not enough for a man approaching his seventieth year, Planck also undertook to defend a causal, realistic philosophy of physics in the new age of quantum mechanics.

The problems Planck dealt with in the Weimar era were minor compared to those he confronted after Hitler took power in 1933. As secretary of the Academy and president of the Kaiser-Wilhelm-Gesellschaft, Planck could not avoid dealing with the government. Convinced at first that the Nazi excesses could not last, and that Hitler's call for national revival and unity had much to be said in its favour, Planck was ready to accommodate to the demands of the new regime. At 75 he would have been relieved to give up the burdens of his offices, but he felt responsible to those who counted on his help and trusted his "immensely noble purity".

Planck's hopes for quick return to civilized behaviour in Germany were quickly dashed when his Jewish colleagues were forced out of their positions and many of them left the country. Still reluctant to make public statements, Planck did not protest against the Academy's offensive response to Einstein's resignation. He went on advising his younger colleagues to have patience and stay at their posts. Nevertheless, it was Planck who went directly to Hitler in May 1933 and tried to convince him that his policies were destroying German science. Hitler would make no distinctions, and became so enraged that Planck could only "fall silent and take my leave". His efforts to salvage what could be salvaged, to block (behind the scenes) the irrational behaviour of his fellow Nobel laureates Philipp Lenard and Johannes Stark, both enthusiastic admirers of Hitler since 1924, did have some positive results. More and more frequently, however, he had to compromise his principles and accept the unacceptable.

During the Second World War, now an octogenarian but free of offices at last, Planck continued his public lectures on science and philosophy, using them as a platform for defending the values of science and civilization and for implicit criticisms of the regime. The disasters of the war, and especially the arrest and execution of his son Erwin for complicity in the 1944 attempt to assassinate Hitler, finally broke Planck's spirit. He died in 1947, just short of his ninetieth birthday.

Heilbron's thoughtful book makes it clear that there were no easy answers to the dilemmas Planck faced. Einstein did not forgive Planck for his public silence in 1933, but he had never lived with Planck's responsibilities. One can speculate on what might have happened if Planck had made a different series of decisions, but speculation is only speculation. Planck acted in accordance with his principles and his conscience, though the world crumbling around him had little place for either. Heilbron writes in his preface that Planck's "life is a lesson", but each reader must decide for himself what that lesson is. ☐

Martin J. Klein is Higgins Professor of the History of Physics at Yale University, Box 2036, New Haven, Connecticut 06520, USA.

Matter of utmost gravity

Joseph H. Taylor

Was Einstein Right? Putting General Relativity to the Test. By Clifford M. Will. *Basic Books: 1986. Pp.253. $18.95.*

IN THE centenary year of Albert Einstein's birth, 1979, a number of books appeared which described the man and his science, especially general relativity. Seven years later, Clifford Will has added another to the list — a readable volume detailing the renaissance that has taken place in gravitational physics since a few years after Einstein's death in 1955. Writing for a general audience, Will describes (sometimes with first-hand knowledge, and always with flair) developments of the past quarter-century which provide the underpinnings of our understanding of the gravitational interaction in nature and its apparently inextricable connection with the geometry of the Universe.

The 1960s, heyday of radio astronomy, dawn of the space age and beginning of the era of atomic clocks, saw for the first time the development of experimental techniques well suited to the high-precision measurement of tiny differences between the predictions of newtonian gravitation theory, general relativity and other "currently viable alternatives". Will tells us of these developments, taking a topical approach. He starts with the reincarnated "Eötvös experiments" carried out in Princeton and the Soviet Union, which verified the equivalence principle to very high precision. Gravitational red-shift tests are described next, with emphasis on the Pound–Rebka–Snider experiments using gamma rays in the Jefferson Tower at Harvard, and the "flying atomic clocks" experiments carried out by Robert Vessot and others a few years later. Together, these tests showed conclusively that a correct theory of gravity must be a metric theory, embodying the concepts of curved space-time.

Subsequent chapters deal with modern experiments measuring the bending of radio waves by the Sun's gravity, the perihelion shift of Mercury by radar ranging techniques, the slowed propagation of signals in the solar gravitational field, extension of the equivalence principle to include the self-gravitational binding energies of the Earth and Moon, the constancy of the gravitational "constant" G, and the loss of orbital energy in the binary pulsar system by gravitational radiation. Throughout the narrative, Will sprinkles a trail of personalities and anecdotes. Robert Dicke's name appears frequently: first as the author of a 1961 article in *Scientific American* which helped to convert the teenaged Clifford Will from a career in architecture, or maybe genetics, to one in physics; later as quintessential experimenter, and co-author of a theory of gravitation to rival Einstein's; later still as a passenger on an airplane bound for Missoula, Montana, on which he was informed by Kenneth Nordvedt, then a total stranger, that the Brans–Dicke theory appeared to be inconsistent with the equivalence principle for strongly self-gravitating bodies.

Irwin Shapiro's discovery of the gravitational propagation delay for electromagnetic waves makes a good story; it includes an image of Shapiro making an eleventh-hour pitch — via long-distance telephone connection from Cambridge, Massachusetts, to the Jet Propulsion Laboratory in California, complete with locally shown viewgraphs, NASA style — pleading for inclusion of the time-delay experiment on the forthcoming Viking mission to Mars. The story of the discovery of the Hulse–Taylor binary pulsar is here, too, complete with graphical sketches reproduced from Russell Hulse's laboratory notebook and with the correct inference that graduate student Hulse and assistant professor Taylor were thoroughly astonished by the potential implications of what they had found.

Lincoln Laboratory. Reproduced from *Was Einstein Right?*

Radar ranger — the Millstone Hill antenna, northern Massachusetts, used by researchers at the Massachusetts Institute of Technology from 1959 for radar echo observations of the planets.

Will has succeeded in what he set out to do, namely to write an interesting and accessible account of recent progress in one field of physics. ☐

Joseph H. Taylor is James S. McDonnell Distinguished University Professor of Physics at Princeton University, Princeton, New Jersey 08544, USA.

EXPLORING ENERGY . . .

LEARNING ABOUT ENERGY

by **David J. Rose**

A wide-ranging text for senior and graduate students that weaves technical and scientific information into an integrated current account of the field, including energy in its social and economic contexts, environmental consequences of energy use, energy conservation, fossil fuel resources and reserves, coal, nuclear power, solar power, and global electrification. A volume in the series Modern Perspectives in Energy.

0-306-42124-0/528 pp/ill./1986/$59.50
($71.40 outside US & Canada)
text adoption price on orders of six or more copies: $39.50

NEW FROM PLENUM TRADE

EINSTEIN'S DREAM

The Search for a Unified Theory of the Universe

by **Barry Parker**

"Completely non-mathematical, and beautifully written. Every page is a pleasure."

—**Isaac Asimov**

Award-winning writer Barry Parker depicts the men and their theories in the search for ultimate answers in modern physics.

0-306-42343-X/298 pp./cloth/$18.95
($22.73 outside US & Canada)

SCIENCE TRIVIA

From Anteaters to Zeppelins

by **Charles J. Cazeau**

"A work of reference and a delightful bedside book."
—**Glyn Daniel**, *University of Cambridge*

Let's Explore columnist to 81 Gannett newspapers presents questions that will pique the curiosities of the Ph.D. as well as the grade-schooler.

0-306-42353-7/274 pp./cloth/$17.95
($21.54 outside US & Canada)

THE LANGUAGE OF SCIENCE
Plenum
PUBLISHING CORPORATION

Plenum Publishing Corporation
233 Spring Street
New York, N.Y. 10013-1578
In the United Kingdom:
88/90 Middlesex Street
London E1 7EZ, England

Reader Service No.1

Doctor, doctor...

Nicholas A. Rupke

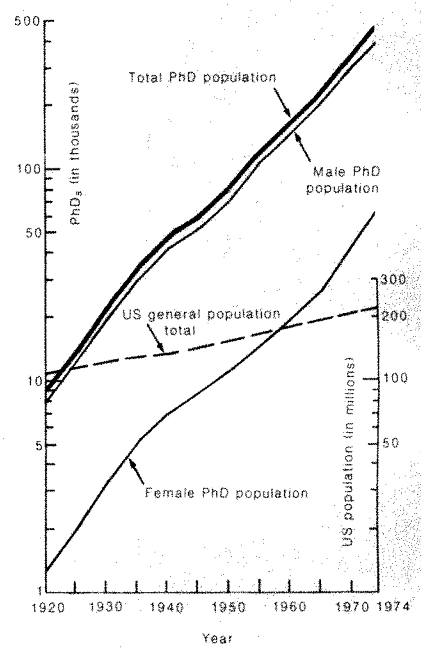
Frontiers of Learning: The Ph.D. Octopus. By Pjotr Hesseling. *Foris Publications:* 1986. Pp. 204. Hbk £27.95, $44; pbk £17.95, $27.*

"CANDIDATES must have a PhD . . .". This stipulation has become a standard phrase in the job advertisement columns of science magazines. The PhD, it would appear, has come to represent something of a passport to academic jobs and a union card for the would-be industrial scientist. Employers routinely use the qualification as a preliminary screening device to whittle down large numbers of applicants, often for jobs in which research training has little relevance. Although a result has been a steep rise in the number of people holding a doctorate, especially in the United States, relatively little is known about the PhD phenomenon. What is the function of the degree, its intrinsic worth, as the crowning feature of higher education? What fruits do we expect it to yield? What precisely should it represent? Is the recent multiplication of doctoral degrees a sign of academic health or a craze that is devaluing them?

It is questions such as these that Pjotr Hesseling, professor of business economics at the Erasmus University of Rotterdam, addresses in this book. He uses a comparative approach, looking at the history of the PhD in different countries and cultures. In its modern form the PhD is a German invention, the foundation of the University of Berlin in 1810 marking the beginning of higher education centred on the research imperative. The doctorate became less of an ornamental initiation rite and more of a serious research accomplishment, representing an independent and original contribution to knowledge, embodied in a book-length dissertation. From Germany the PhD moved to the United States (first award in 1861), Britain (1917) and more recently to Japan (1953), gradually penetrating the higher education system of each country.

In the course of its global advance, the PhD has been adapted to various indigenous cultural traditions. The European doctorate is to some extent still a shibboleth of originality and cultural authority (although the precise appreciation varies between disciplines and countries); the American doctorate is primarily used for pre-career training and as a certificate for entering the academic and other professions; the Asian doctorate, which Hesseling has studied mainly in Indonesia and Japan, is still in a formative phase, being made to fit old traditions of scholarship and wisdom on the one hand, and new, pragmatic considerations of state-managed science policy on the other. Despite these differences, the American graduate school, envied for its productive efficiency, is increasingly becoming the model for the rest of the world.

Hesseling regrets the "Americanization" of the PhD, which is turning it into market-orientated training during which a candidate will often be assigned a narrowly defined topic within a pre-existing research programme. The PhD, he believes, is becoming too much part of government-directed science policy and of university administration preoccupied with production costs. Quality should be what matters most, but this feature is too elusive to be regulated to any great extent. Originality in research, although the outcome will be less certain, should have priority over completion rates — the reverse of what a recent OECD report, *The Role and Function of Universities: Post-graduate Education in the 1980s*, says on the subject.



PhD explosion — the rise in the number of living holders of doctorates earned in the United States between 1920 and 1974.

Redrawn from *Frontiers of Learning*

Hesseling does not believe that the doctoral dissertation is a necessary complement to the education of all postgraduates. To him, the PhD should be part of a special "cultural mission" which includes testing received knowledge, overcoming boundaries between disciplines, institutions and nations, and striving for such ideals as disinterestedness and communalism. A doctorate should explore, as the title of his book states, "The Frontiers of Learning" and not be an overspecialized tentacle of "The Ph.D. Octopus" (a sub-title borrowed from the Harvard psychologist William James).

---

*PO Box 509, 3300 AM Dordrecht, The Netherlands, and PO Box 5904, Providence, Rhode Island 02903, USA.

To achieve this end, Hesseling advocates a European doctorate through the revitalization of the traditional academic peregrination which made use of the advantages of different systems A code of practice should be adopted by the main European universities with, among other things, the abolition of fixed completion periods for doctoral projects, emphasis on interdisciplinary work, freedom for the candidate to select both topic and supervisors (the latter from more than one university), use of external examiners (already the practice in Britain) and publication of the thesis in a major language

One can readily sympathize with these ideals But given the increasingly tight and competitive marketplace, and given also the domination of the American system, the need to hold a doctorate as well as a bachelor's and a master's degree is greater than ever It would therefore seem unrealistic to revert to a system of higher education in which the PhD is for the relatively few Furthermore, Hesseling's ideals are to a large extent accommodated, especially in the Anglo-American world, by the network of post-doctoral fellowships which provide an opportunity to elaborate and publish doctoral theses Many of these fellowships are held in institutes for advanced study with their interdisciplinary and international atmosphere The Princeton Institute has been replicated, *mutatis mutandis,* not only in the United States itself but also in Germany, The Netherlands and Scotland In England, several of the Oxbridge colleges give hospitality to advanced researchers in different disciplines

Closer examination of these and related questions is obviously needed Hesseling's book is seriously flawed stylistically — there are innumerable mistakes of English usage and unnecessary repetitions of material, and many passages are exasperatingly vague But he has brought a commendable breadth of approach to his subject, and has provided much food for thought for anyone interested in higher education □

Nicholas A Rupke, Wolfson College, Oxford, OX2 6UD, UK, is a geologist and historian of science and medicine

---

Quantum question

Philip Pearle

The Shaky Game: Einstein, Realism, and the Quantum Theory. By Arthur Fine
University of Chicago Press 1986 Pp 179 $25, £21 25

EVEN such a loving biographer of Einstein and his physics as Abraham Pais seems embarrassed by Einstein's position on quantum theory In his book '*Subtle is the Lord*', Pais quotes Einstein — "I feel that the real joke that the eternal inventor of enigmas has presented us with has absolutely not been understood as yet" — and then remarks, "It is believed by nearly all of us that the joke was understood soon after 1925 " Given this belief, and the spectacular successes of quantum theory which nurture it, it is perhaps not surprising that sympathetic and careful analyses of Einstein's attitude have been lacking

This lack is quite satisfactorily addressed in five (the second to sixth) of the nine essays which make up this book The essays, most of which have been previously published, are clear, well-reasoned and appropriate settings for the gems they are studded with, quotes from the Einstein archives at Princeton The sixth essay, "Einstein's Realism", forms a bridge to a rather different piece of country, three provocative articles on realism which reject both it and anti-realism in favour of the author's own position, which he calls the "Natural Ontological Attitude" (NOA)

The second essay, which follows an introductory chapter, attacks the often-promulgated view that Einstein's position on quantum theory represented a "kind of scientific senility" After giving ample evidence that Einstein's way of thinking supported Poincaré's admiring assessment of " the facility with which he adapts himself to new concepts He has not remained attached to classical concepts ", Fine refreshingly concludes that " it was Bohr not Einstein who felt bound by the classical concepts", and "in the end Einstein was more radical in his thinking "

The third essay details Einstein's route to the conviction that quantum theory is not " in principle capable of producing a complete description of an individual physical system", and culminates in a discussion of the famous Einstein–Podolsky–Rosen paper For those of us who have secretly never felt fully at ease with the tortuous arguments in that paper, and who have wondered how that master of clarity, Einstein, could have done this to us, Fine brings welcome relief He shows just how complicated the argument of the paper is, presents evidence that Podolsky wrote it without revision by Einstein and analyses a simpler argument that Einstein sent to Schrodinger in a letter The fifth essay also concerns letters from Einstein which stimulated Schrodinger while he was writing his famous paper on the cat paradox This paper showed that quantum theory can predict a state for a cat that is a superposition of cat-alive plus cat-dead, an unrealistic situation for a single cat, thus supporting Einstein's statistical interpretation of quantum theory " the description not of a single system but of an ensemble of systems" How delightful to learn that Einstein himself initiated this paradox, with an example of unstable gunpowder in a quantum state of unexploded plus exploded[1]

I found essay four on Einstein's statistical interpretation to be rather incomplete Perhaps because Einstein felt that the de Broglie–Bohm "hidden variable" model is not fundamental enough, no discussion is given of this or the models of Wiener–Siegal or Nelson, which many feel are nice examples of how a statistical interpretation may function Instead Fine presents some of his own work, a refutation of a model he calls "The Standard Answer" (a sharper earlier refutation due to Kochen and Specker is not mentioned) and a possibly satisfactory "Prism Model" whose similarity to the so-called "loophole" models and vulnerability to experimental refutation is not mentioned (until the last essay)

Einstein's realism is analysed next, understood as a faith, and dubbed "Motivational Realism" Fine shows that Einstein, while driven by a passionate realism, drew back from making logical assertions connecting physics to that reality Apparently philosophers are not so judiciously reticent, and Fine rebuts both realist and antirealist positions Instead he opts for NOA, a minimalist position, that accepts the truth of what is useful (for example the electron as a set of working ideas and experimental results) without further interpretation

The NOA may be adequate for a spectator of the game of physics, but I do not see it as appropriate for a player Fine tells us that NOA makes it easy to adopt "paradigm shifts", but that is passive behaviour Physics, the active game of creating beautiful pictures with predictive power, was played so well by Einstein not least because of his realist faith

Fine rather snidely compares the realist faith to religious faith In the same vein I would add philosopher's faith (in unexamined axiomatic concepts), for it seems each philosopher is refuted by another on such "religious" grounds There is nothing wrong with faith all physicists act on more knowledge than they can prove But in matters of faith, certainly evidence counts The creative successes of realists over the centuries justify their leap of faith more than that of most others And, although one might attribute the success of quantum theory in the past 60 years to nonrealists, in practice they deal with the reality of the concepts they use the way Bohr dealt with a horseshoe He put one on his cottage door, and said, "No, I don't believe in it, but I understand the good luck comes even if you don't believe" □

Philip Pearle is a Professor in the Department of Physics, Hamilton College, Clinton, New York 13323, USA

Making waves on Shelter Island

Ya B Zeldovich

Shelter Island II. Proceedings of the 1983 Shelter Island Conference on Quantum Field Theory and the Fundamental Problems of Physics. Edited by Roman Jackiw, Nicola N Khuri, Steven Weinberg and Edward Witten *MIT Press 1985 Pp 369 $37 50, £37 50*

The proceedings contained in this book are best viewed against the background of the evolution of physics in the twentieth century and, more specifically, against events at the previous Shelter Island conference of 1947 That meeting was organized by Robert Oppenheimer and was attended by many talented scientists who had returned to fundamental research after working on war projects

The intermittent character of scientific progress is now well-recognized, bursts of fermentation of new ideas are punctuated by steady elucidation of all the consequences A Russian poet Tjutchev, practically unknown abroad, said "Blessed are those who come into this world at its fateful moments", for physicists the remark can be paraphrased as "Lucky are those who attend conferences when fresh ideas are in the air"

One such period was between 1905 and 1928, which witnessed the advent of special and general relativity, quantum mechanics, the quantum theory of fields and the theory of antiparticles In the 1930s, development slowed down the refinement of theory was plagued by infinities A theory of nuclear forces was proposed, with heavy particles — mesons — being an essential part of it Particles with approximately the appropriate mass were found in cosmic rays, but the euphoria turned to desperation when they were found not to fit the theory Many eminent physicists believed that a new conceptual framework would be necessary, something as revolutionary as relativity and quantum theory as Bohr is reputed to have said "This theory is not crazy enough to be true"

But the younger generation of the time (and many of the old) were more sober Shelter Island I gave impetus to the work of Schwinger, Feynman and Bethe, whose renormalization programme brilliantly explained the Lamb–Rutherford level shift in the hydrogen atom and the anomalous magnetic moment of the electron Marshak went on to consider the possibility of two mesons the muon as a heavy electron and the pion giving nuclear forces For several decades physics was dominated by ingenious applications of principles formulated before 1930, even gauge theory and quarks fall in this category It was not until the mid-1970s and 1980s that a further wave of ideas arose, and it is these new ideas that are the subject of this volume

In a nutshell, the goal of the old theory was this given the existence and characteristics of particles and fields, how could one move on to calculate all processes, bound states and so on? The new theory has a more ambitious programme It aims at explaining *why* all known particles exist — and determining their masses and charges The same theory also predicts the existence of many as yet unknown particles and fields Very strange — even "crazy" — ideas are needed!

An excellent overall view of how things stand is given in the first paper in the book, where Murray Gell-Mann provides a coherent yet witty account of how theory has developed between the two Shelter Island conferences Here, too, Gell-Mann summarizes the topics treated in detail in the following articles — supersymmetry, Kaluza–Klein multidimensional theories, superstrings, and the new and fruitful symbiosis between particle physics and astrophysics

To a reader not well acquainted with particle physics, I should also add mention of two events of previous history the introduction of Planck units of mass, length and time in 1899, and the appearance of Dirac's famous paper on large numbers in physics the ratio of electrostatic to gravitational interaction, 10^{39} in the atom, for example The modern answer is quite different from that proposed by Dirac, but that does not invalidate the importance of the question he raised

Characteristic of the new wave of ideas is Steven Weinberg's paper, "Calculation of Fine Structure Constants", in which he discusses the now popular idea of higher dimensional space-time manifolds Time and three spatial coordinates describe the usual space-time, while the extra spatial coordinates are assumed to be "compactified" to dimensions of the order of 10^{-32}cm They are not observable as coordinates, but their deformations manifest themselves as fields and particles As shown in Witten's contribution later in the book, the structure of these extra dimensions may explain the spectrum of fermions — the existence of three, and only three, families of leptons and quarks

This is a revival of a theory, initiated in the 1920s by Kaluza and Klein, to unify gravity and electromagnetism For a long time Einstein's own search for a unified theory appeared to be simplistic and to be contradicted by the proliferation of fields and particles found by experimenters Now, however, most theorists are convinced that the time is ripe for brainstorming in this direction but with more dimensions and a peculiar symmetry (supersymmetry) between bosons and fermions

A new branch of physics, as compared with the 1947 conference, is the interdisciplinary study of the very early Universe Then, because of the negative pressure of a scalar field, one had gravitational repulsion instead of attraction and the radius of the Universe grew exponentially Both Guth and Linde discuss the topic (Incidentally, Frenkel and Landau from the USSR were invited to attend Shelter Island I, but neither of them went, Linde, of the Lebedev Institute in Moscow, did attend in 1983, and contributed The ratio 1 0 is a great improvement!) In the 1960s, the big bang theory of the origin of the Universe was supported by the discovery of background microwave radiation in the Universe, and by the theory of nucleosynthesis occurring at temperatures from 3×10^{10} to 3×10^{8}K These energies are within the range of our particle accelerators, results from which have lent further weight to the big bang theory But in attempting to go back to an even earlier stage one enters unexplored territory, and the Universe itself has to play the role of the poor man's accelerator Here the astrophysicist's work is akin to that of the palaeontologist, in which events in a past era must be interpreted from their traces observed today For example, in my opinion (shared by many others, including Gell-Mann) the very charge asymmetry of the Universe, that is, the absence of bulk antimatter, is proof of the instability of the proton, which Earth-bound experimentalists have searched for in vain

Everybody working on or interested in the fundamental problems of physics should read *Shelter Island II* One needs a good grasp of quantum electrodynamics and some special mathematics to appreciate the contents fully, but even those without such knowledge will gain an awareness of contemporary developments in theoretical physics The science-appeal emanating from the book is very strong, I suspect that many physicists will read it initially without understanding everything, and then learn more fundamentals and return to it

Although the conference was held in 1983, the book has not dated very much, the various authors remain leading authorities on their subjects and are now, three years later, still developing the ideas discussed in their contributions Altogether one is grateful to the organizers of the conference and to the editors for bringing together such influential people and for collecting their papers into this excellent book The publishers might consider issuing a cheaper paperback edition so that its contents would be more widely available, especially to students ☐

Ya B Zeldovich is at the Institute for Physical Problems, USSR Academy of Sciences, Ul Kosygina 2, Moscow 117973 GSP-1, USSR

nature

NATURE VOL 324 20 NOVEMBER 1986

Other people's nuclear weapons

Week-end diplomacy seems to have won Britain a prolongation of its role as a nuclear power, but the time has come to ask when and how this will be attenuated.

THE British government's case for maintaining nuclear forces is far from simple and far from clear. Historically, the British stumbled into making nuclear weapons because, in the late 1940s, it seemed to the then Labour government that this was the natural course to follow in a substantial and technologically advanced country. Over the years, circumstances have changed. The range of Britain's international interests has shrunk to an extent that could not have been foreseen forty years ago, while the country's industrial economy is no longer large enough to sustain the independent development and construction of strategic delivery systems.

The obvious parallel with the present time is that in the early 1960s, when the then Prime Minister, Mr Harold Macmillan (now Lord Stockton), persuaded President John F Kennedy that, if the air-launched nuclear bomb called Skybolt was not available (because it did not function as designed), the British government should be allowed to purchase Polaris submarines instead. But why? In one sense, it is merely right and proper that a government which is a member of an alliance such as the North Atlantic Treaty Organisation (NATO) should make an equitable contribution to its strategic forces, indeed, the British Polaris submarines are operated within NATO guidelines, but with the understanding that they might be withdrawn, and operated independently, if some vital national interest were involved.

But what could those circumstances be? Is not the United States fully committed to the defence of Europe not merely through the formal treaty on which NATO rests but also by the presence both of troops and of strategic missiles on European soil? The simple answer is that these assurances have been insufficient for a succession of British governments, all of which have recognized that circumstances could arise when the United States might shrink from using its strategic nuclear forces, risking retaliation in the process, for the sake of some ill-understood European interest. These dark suspicions have been strengthened by the news from Reykjavik last month.

Negotiations

Whether Britain and France will be able indefinitely to continue in that frame of mind is another matter. Everything will depend on what happens in the negotiations on strategic arms still continuing at Geneva. After Reykjavik, and the recognition by the two most powerful people in the world that even they cannot magic out of thin air lasting agreements on difficult questions in a dozen hours of conversation, it is inevitable that strategic arms control will be regarded as a matter for the long haul, the bilateral negotiations at Geneva will (with luck) be prolonged into the next decade, and more attention will be paid to other aspects of the relationship between the major powers. But the time will come when, again with luck, there will be another proposal that the balance between both strategic missiles and missiles of intermediate range should be struck at a lower level than at present. At some stage, the point will (with luck, again) arrive at which the British and French nuclear forces stand out conspicuously. Then the superpowers will be pressing to see these independent nuclear forces cut back to size. The United States, having helped Britain to build Trident submarines, will be wanting to see them scuttled. That will be an awkward conversation for some future weekend at Camp David.

The British position on this forseen dilemma is, fortunately, explicit. In the first place, the submarine missiles are regarded as strategic weapons, contributing now to the strategic deterrent of NATO as a whole but also intended as a means by which Britain might threaten even a more powerful potential adversary with an unacceptable degree of retaliatory damage. Theoretically, the French regard for nuclear weapons is very much the same, they might not be sufficient to influence events on the global stage, but they would surely 'tear an arm' from an attacker, and give him pause. But the British government has also gone so far as to consider when it would be proper to think of limiting its nuclear forces, saying that this would be acceptable as part of a more general agreement for the substantial reduction of nuclear forces everywhere.

Unfortunately, this undertaking is not nearly as clear as it might be. The British government argues that, because its weapons are strategic, their number should not be affected by an agreement on the numbers of superpower nuclear weapons deployed in or against Europe, which makes sense. But the prospect that there may be very substantial reductions of European nuclear forces can only strengthen the British government's unspoken fears that Europe may one day be left to its own devices, whence in part Mrs Thatcher's visit to Washington at the weekend. The snag is that the British position is vague on what would happen if there were also a substantial agreement on strategic forces might be held to provide a licence for conventional wars in Europe, in which West European states would be disadvantaged. So British governments will be tempted to hang on to their nuclear weapons through thick and thin.

The lessons to be learned from these tortuous arguments are several. First, Britain's susceptibilities about its nuclear weapons mirror those of other European states about other aspects of their arrangements for defence. British nuclear weapons would indeed be an impediment to sudden and drastic reductions of nuclear arms along the lines fancied at Reykjavik, but so would be the political protestations of other European states, from which it follows that the superpowers must aim in their negotiations at a balanced package of arms control agreements covering not merely strategic and intermediate nuclear missiles but conventional forces and even the temper of East–West relations. On that score, the Europeans have a case. But the United States also has a right to expect that the British and French nuclear weapons will not be impediments to reasonable arms agreements that may in future be reached with the Soviet Union. Indeed, it would have been natural that President Reagan should have broached the question with Mrs Thatcher at the weekend, while confirming that Trident should succeed Polaris. And, given governments' temptations to hang on to weapons in the face of arguments to the contrary, there is the strongest case for asking that the British government should now be prepared publicly to define the circumstances in which it will let go. France, which is a harder case, should be persuaded also to follow suit.

Carnegie's misdiagnosis

A new study of undergraduate education misses an opportunity to tell what is really wrong.

US ACADEMICS and their followers, in one of their periodic fits of introspection, have been giving pious attention to a report from the Carnegie Foundation for the Advancement of Teaching on the state of higher education and, in particular, to its judgement that, "driven by careerism and overshadowed by graduate and professional education, many of the nation's colleges and universities are more successful in credentialing (*sic*) than in providing a quality education for their students". There will be few who disagree with that, but the foundation's study, *College — the undergraduate experience in America*, by Mr Ernest Boyer, is most notable for what it has left unsaid.

Carnegie is, naturally, right to be troubled by the lack of individual attention given to undergraduates and to sympathize with them in the hard choice of a course of study. Then few will quarrel with Boyer's plea for better personal advice to students, especially during the difficult transition from high-school to college. And nobody will argue against the case for higher standards of literacy. But the argument hangs too much on the subjective experiences of undergraduates and teachers, as reported and, often touchingly, recorded. This preoccupation with tensions between "individual interests and shared concerns" means that some of the urgent questions about the future of higher education in the United States are mis-stated.

One is the pattern of the curriculum. US universities have always required less specialization than have universities elsewhere, and deserve credit for it. Their accessibility is unmatched anywhere in the world. But the worry now is that things have gone too far, and that US undergraduates have too much latitude to choose from a curriculum increasingly resembling a supermarket for credit courses, with students free to experiment as they please until the required number of credits is reached. The Carnegie report rightly raises this important issue, with which William J. Bennett, the US Secretary of Education, startled some of those at Harvard University's 350th birthday celebrations earlier this year, taking Harvard as his inappropriate example. Carnegie's remedy for the laundry-list approach to education is to call for a core curriculum. The need for some high-level coordination cannot be denied. But if Carnegie wishes to be taken seriously, it must devise a more persuasive core than the amalgam now put forward, an "integrated core" that would include language, art, history, institutions, nature ("the ecology of the planet"), work ("the value of vocation") and identity ("the search for meaning"). All this, but no general science course?

The Carnegie study also draws attention to the conflicts that can arise in research universities between the demands of teaching and of research. The story is familiar: because academic prowess is measured largely by published research papers, faculty members devote their best efforts to excellence in research, sometimes at the expense of teaching. Boyer's suggestion that research universities should establish a new rank of "distinguished teaching professor", extending status and salary incentives to outstanding teachers, might serve well in some colleges and universities, but ignores the frequency with which excellence in teaching and in research go together, hand in glove. In the end, the research universities will have to deal with this internal conflict internally, by requiring that faculty members are not so distracted by external interests and commitments as to impede their discharge of their academic responsibilities, to students as well as to research. This goal will not be easily won, but the research universities eager that their reputations should survive will give it more attention than has been the habit in recent years. The professorial status of good teachers is a palliative only.

The Carnegie Foundation would also have done education a greater service if it had tackled some of the problems of higher education that are beyond the control of the universities themselves. The most immediate threat is the accumulated effect of the dereliction of university research facilities and cutbacks in support for students. Many of the institutions that have served the United States well in the past, the less well-known universities that dominate the Boyer's survey in particular, are now ominously stretched. Government officials are fond of pointing out that support for basic research at universities has increased in recent years, which is true. But the increase followed a long decline during the 1970s that has not yet been made good. Decaying facilities and equipment are depressingly common on all but the most successful campuses. Federal support for university research facilities dropped from $404 million in 1966 to a low point of just $37 million in 1981 (1981 prices). Support for graduate students in the sciences has fallen by about a half since 1969. The efforts of the National Science Foundation to promote excellence in science education, all but abolished by President Reagan at the start of his first term of office, are improving, but slowly. Small wonder that the proportion of bachelor's degrees in physical and biological sciences has fallen since the late 1970s, with even engineering starting to tail off in 1985. Yet fewer undergraduates now means fewer faculty members in ten years' time, less research and, probably, less teaching. Unless the federal government decides to act, these are the circumstances that will condition some future Carnegie report on the parlous condition of higher education in the United States at the end of the decade. □

Fire without smoke

Two retracted papers do not make a scandal, but too much reticence about the circumstances may.

DANA-Farber Cancer Institute has a problem (see page 197). It has announced that two papers describing a T-cell activating substance called interleukin-4A are being retracted from the journals in which they were published. This is certainly not the first time that scientific papers have been withdrawn, and it will not be the last. It is very much to the credit of Dana-Farber that the retractions were issued as soon as the errors in the data were discovered.

Where Dana-Farber's problem lies is in the way it described — or more properly failed to describe — the reasons for the retraction. Obviously there are many reasons why papers may have to be retracted. Cell culture contamination, reagent impurities, inadvertent miscoding of results can all be discovered, long after the fact, even in the most careful and thorough laboratories. But Dana-Farber has not invoked any of these readily believable explanations, choosing to make the blunt statement that an investigation is under way, and that retraction letters have gone out. This approach can only leave a cloud of suspicion hanging over the institute. What sort of investigation is it that can immediately conclude that the data upon which a paper is based are incorrect and not reproducible without saying how that fact is known?

Unhappily, there have been all too many recent instances in which falsification of data has brought laboratories into trouble. By declining to comment about its investigation, Dana-Farber, perhaps unwittingly, merely encourages speculation. Officials at Dana-Farber will correctly say that it is unfair to make allegations that could have a profound influence on a scientist's career before they are completely investigated. But silence is not the answer. Surely it is possible to say what first aroused suspicions about the data without jeopardizing future investigation? Surely retractions with no explanation will focus interest on the case longer than a simple statement about the circumstances as they are at present known? The scientific community should be credited with the good sense to understand the distinction between an investigation and a final judgement. □

Research publishing

Interleukin disappears from the literature

Boston

Two papers from Dr Ellis Reinherz's laboratory at Dana–Farber Cancer Institute that announced the discovery and properties of a T-cell activating substance termed interleukin-4A(IL-4A) have been retracted by the authors The reason given for the retraction is that the data are incorrect and are not reproducible The authors, in other words, no longer claim to have proof of the existence of IL-4A originally hailed as the chemical messenger that amplifies the body's immune response by turning on resting T cell's, through its interaction with the T11 protein at the surface of the T cell

The papers being retracted were published in *Science* (**231**, 1118, 1986 by Claudio Milanese, Neil E Richardson and Ellis L Reinherz) and the *Journal of Experimental Medicine* (**163**, 1583, 1986 by Claudio Milanese, Robert E Siliciano, Reinhold E Schmidt, Jerome Ritz, Neil F Richardson and Ellis L Reinherz) Letters of retraction are being sent to both journals

An *ad hoc* committee of scientists from both within and outside Dana–Farber is investigating why the papers need to be retracted Both Reinherz and David Kiszkiss, director of research at the institute, refused to comment on the reasons the data became suspect, how the data were proved incorrect, or why a retraction of the articles was necessary even before an investigation was completed It is possible that litigation may result, according to Reinherz, but a final decision will be up to Dana–Farber and its parent Harvard Medical School or perhaps the National Institutes of Heath (NIH), Kiszkiss says

Kiszkiss refused to comment on the process of data review at Dana–Farber or in the Reinherz laboratory Dr Sheldon Wolff of the New England Medical Center, says any person in a laboratory understands there will be efforts to reproduce all published data There is no way the principal scientist can be expected to stand at the elbow of everyone working in a laboratory, says Wolff The work of Reinherz is unanimously described as "impeccable" by others in the field

The withdrawal of the IL-4A papers means there is now only one well-studied candidate for a T11 ligand, a molecule termed LFA3 The data supporting the existence and activity of this molecule are so convincing, says Dr Peter Beverley at the Imperial Cancer Research Fund in London, that "a number of people were sceptical of the dramatic announcement" of IL–4A According to Dr Stephen Shaw of NIH, a principal investigator of LFA3, there is much interest in finding the one or more molecules that interact with T11, also called CD2 T11 is a marker on T cells thought to be involved in adhesion of lymphocytes and triggering of T-cell activation IL-4A was described as soluble, of low molecular weight, and responsible for triggering resting T cells LFA3 is a much larger cell-bound molecule responsible for lymphocyte adhesion

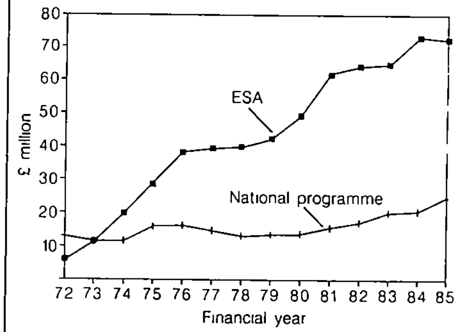
Shaw says he originally accepted the IL-4A data because they came from the Reinherz laboratory and because there is room on the T11 molecule for two different recognition sites He also thought IL-4A could be a soluble fragment of LFA3 Even though there is no longer proof of the molecule termed IL-4A, he says, other ligands for the "popular and trendy" T11 molecule may exist, including (he speculates) a soluble form of LFA3

Elizabeth Collins

British space

Waiting for an increased budget

THE British National Space Centre (BNSC) this week celebrates its first anniversary still awaiting the government's response to its 15-year plan for Britain's presence in space The blueprint, if accepted, will mean a substantial increase in the £108 million a year that Britain now spends on international and domestic space projects



BNSC seeks more funds for space research and a narrowing of the gap between international and domestic projects

Britain is under pressure from the United States and Europe to invest heavily It was the need to respond sensibly to an invitation to participate in the US Space Platform that inspired the creation of BNSC (see *Nature* **318**, 305, 1986)

A ministerial council meeting of the European Space Agency (ESA) in Rome in 1985 generated pressure from Europe The meeting approved in principle a number of new programmes that would require members' contributions to be increased by 70 per cent in the next three years France, for example, which already spends five times as much as Britain, has decided to increase its budget by 40 per cent, bringing the total annual expenditure on space to £700 million

Roy Gibson, director general of BNSC, says "the European scene is going on, whether we are ready or not", and that BNSC must have a positive response from government within the next few weeks if proper preparations are to be made for the ESA ministerial council meeting next June at which the final touches to the ESA programme, some lasting more than a decade, will be made

"Last year we weren't too worried about this deadline in June 1987 now we're getting a big edgy because we need these decisions in order to give us the best bargaining position inside ESA "

The British space plan, he says, will require substantial international investment and the domestic budget must be increased to ensure industrial back-up to exploit the new technologies Much of the money for Britain's contribution to the space platform project, Gibson believes, must come from the government Although the private sector may eventually benefit, it would find it hard to justify a commercial investment at this stage "I think it would be very wrong to tell government that it's going to make a lot of money straight away "

BNSC wants the government to commit itself to plans that will span more than a decade and to allocate money at least three years ahead to allow industry to respond and plan accordingly

To that end, the BNSC plan details the principal areas in which the United Kingdom must develop These are space science, remote sensing, microgravity, telecommunications (with particular emphasis on mobile communications on land, air and sea), space station, ground/test/reception facilities, data processing and enabling technologies

The present budget is culled from the space funds of the Department of Trade and Industry, the Ministry of Defence, the Natural Environment Research Council and the Science and Engineering Research Council The 280 members of staff are spread between London, the Rutherford Appleton Laboratory in Oxfordshire and Farnborough in Hampshire The next stage would be to create a permanent role for BNSC and perhaps a permanent home

Bill Johnstone

Soviet accusation

Plagiarism charges levelled

Two leading members of the Academy of Sciences of the Byelorussian SSR have accused a US professor, Dr Hollis B Chen, of plagiarism Writing in the Russian weekly *Literaturnaya Gazeta,* Dr Mikolaj Barysevic, president of the Byelorussian Academy, and Dr Barys Sciapanau, a three-times State Prize winner, have accused Chen of pirating chapters 3–6 of his textbook *Theory of Electromagnetic Waves — a Coordinate-Free Approach* (McGraw Hill, 1985) This material, they claim, was lifted without acknowledgement from two monographs by Dr Fiodar Fiodarau, son of the Byelorussian writer Janka Maur Fiodarau is a leading light of the Byelorussian Academy, and is secretary of its section of physical, mathematical and technical sciences

According to Barysevic and Sciapanau, the "coordinate-free" method was pioneered by Fiodarau more than 30 years ago On the basis of this "fruitful approach", they say, he has published six monographs and "hundreds of scientific articles" Chen's work, they allege, is based on two of three works, the monograph *Optics of Anisotropic Media* (Minsk, 1958) and the book *Theory of Gyrotropy* (Minsk, 1976) Both works are fairly obscure, having been published, the critics admit, "a fairly long time ago, in a small edition, and from an obscure scientific press (*Navuka i Technika,* Minsk)

Their accusations may be summarized under four points

- Chen, they say, claims on the blurb that his method is new, and that the book "for the first time treats the essence of the subject fully on a coordinate-free basis",
- Chen's bibliography makes no reference to Fiodarau in its more than 200 items,
- Chapters 3–6 of Chen's book are almost completely copies from Fiodarau, and, in particular, the diagrams have been lifted unchanged,
- Chen's book, "contrary to academic tradition", does not give references, within the body of the text, to the works mentioned in the footnotes This, say the critics, is to avoid revealing how much material is derived from Fiodarau

The complaint ends by listing the many learned societies to which Chen belongs, in terms that leave no doubt that, in the opinion of the critics, he should be expelled as "one of the most cynical and shameless plagiarists in the history of science"

Examination of the book does not appear to substantiate these allegations Chen's work, a textbook for final-year undergraduate and graduate students, does not claim any novelty of method, although his preface does speak of its being an "alternative approach", a coordinate-free approach, and notes that "the organization of the book is somewhat different from that normally found in books on electromagnetics" His diagrams, in the disputed Chapters 3–6 in particular, seem standard vectorial representations of such phenomena as refraction and reflection But Chen in fact makes little claim to novelty and was not, he said, responsible for the jacket In a statement to *Nature* last week he says that he had never claimed the creation of a new method, and maintains that he had relied heavily on the previous work of others The book is not, he emphasizes, a research monograph but a university text He therefore makes little reference to work published in languages other than English He has no knowledge, he stresses, of Russian

Vera Rich

Chen's statement

"This book is intended as a text for seniors and first-year graduate students, not Research monograph. The development of the book relies heavily on my previous research work on waves in plasma and the published work of the others as stated in the preface. I have never claimed the creation of the new method. In fact, the vectors and dyads were invented in 1881 by Professor J.W. Gibbs of Yale University Unfortunately, I do not read Russian and most of my references were in English. This is because the focus was on developing a strong textbook in electromagnetic waves rather than a comprehensive research monograph". □

Australian universities

Going private pleasantly?

Sydney

Plans for a private university that would break the public monopoly in higher education are attracting strong opposition from academics, in particular university staff associations and teachers' groups The address of the Bond University of Technology is Miami, Surfers Paradise, on Queensland's Gold Coast, one of Australia's most pleasant areas It will open its doors in 1989 to a first intake of 1,000 out of a planned total of 10,000 students, many of whom will be from overseas, particularly from Japan and South-East Asia But the organizers hope that at least half of the students will come from Australia's free government tertiary institutions

The university is the brainchild of businessman Alan Bond, of America's Cup fame Funds for the university will come equally from the Bond Corporation and a Japanese development company represented by Dr Bungo Ishizaki of the Osaka Univeristy of Commerce The university will be run as a non-profit body with operating costs covered entirely by student fees The backers intend to make their investment pay by capitalizing on products and inventions arising from research carried out in the university

High salaries will be offered to attract talented staff and the position of vice-chancellor is being advertised at a salary of A$150,000, substantially higher than that for a similar post at an existing university

Criticisms of the Bond Univeristy include doubts about the standard of its courses and whether or not the range of subjects to be offered warrants the title of university The main reason for opposition seems to be the threat it poses to Australia's egalitarian education system by creating an elite of those who can afford to pay One fear is that, once the institution is established, it may be successful in pressing the government for assistance, and that with the proliferation of such private institutions, existing establishments will be left with an ever smaller cake to cut up Others worry that the institution will become a last resort for those not bright enough to gain admittance elsewhere And there is concern over the destabilizing effect of the high salaries on the rigid salary structure of the government institutions

Supporters feel that the quality of present universities is suffering because of the restrictions that accompany government funding Dr Ian Bassett, a senior lecturer in physics at the Univeristy of Sydney, believes that salary flexibility is vital to attract outstanding people Present arrangements of the Commonwealth Tertiary Education Commission (CTEC) prevent universities from offering higher salaries even if the extra comes from private sources When the Univeristy of Sydney recently gave its vice-chancellor a few thousand dollars extra, the same amount was taken out of the following year's grant And the present rigidity in salaries fails to recognize that it is more difficult to attract talented staff to some disciplines than to others

In defence of its own plan, the Bond Corporation points out that it is in line with the government's policy of encouraging more contact between academics and industry Furthermore, with a rapidly growing population of 370,000, the Gold Coast region is the seventh largest population centre in Australia and the only one of its size without a univeristy The present climate of restraint in government spending is unlikely to give the region another chance to get one say the supporters of the Bond plan

Charles Morgan

Nuclear energy

India joins enrichment club

New Delhi

India has developed the capability to enrich uranium, according to Dr Raja Ramanna, chairman of the Atomic Energy Commission (AEC); he did not reveal the size of the enrichment plant nor the amount of U-235 produced. Ramanna was speaking at a press conference and reacting to US press reports that Pakistan is now able to make bomb-grade uranium.

For the past four years, the Bhabha Atomic Research Centre (BARC) has been operating a gas centrifuge pilot plant at Trombay near Bombay. Its operation and progress have been closely guarded secrets until now. Unlike Pakistan, which relied on secret imports of design and parts for its Kahuta enrichment facility, BARC is believed to have designed the plant itself.

There are rumours that, based on the pilot facility, BARC is setting up a full-scale gas centrifuge plant near Mysore in Karnataka state in South India. But Iyengar, in an interview on All-India Radio, said the enrichment project was part of BARC's efforts to develop expertise in all aspects of nuclear energy and that there was no plan to set up a bigger plant.

India has no need for enriched uranium as its Candu-type reactors are based on natural uranium fuel and heavy-water moderator. Only the two US-built light water reactors (LWR) at Tarapur at present depend on U-235 fuel imported from France. Under an agreement with the United States, US permission is required for operating the reactors with India's own enriched uranium fuel. According to Iyengar, the enriched uranium will be required for nuclear physics experiments.

India is self-reliant in the total nuclear fuel cycle technology — from mining uranium and commissioning power plants to reprocessing spent fuel. Enrichment was one area where it lagged behind. Having developed the gas centrifuge technology, India is now on equal terms with Pakistan, but AEC sources have dismissed speculation that India's enrichment activity is related to a weapons programme.

US news reports that Pakistan is on the verge of building a bomb have undoubtedly caused alarm in India, but the government has said it will not take any hasty decisions about going nuclear. Replying to questions about the possible nuclear threat from India's neighbour, Mr Natwaar Singh, Minister of State for External Affairs, told parliament that India "will keep its powder dry" while at the same time exploring all avenues for preventing a nuclear race in the subcontinent. India, which conducted a nuclear test at Pokhran 12 years ago, has consistently denied having a weapons programme.

Meanwhile, AEC has announced an ambitious plan to build 22 more nuclear power plants in the next 15 years to raise India's nuclear power capacity from the present 1,100 MW to 10,000 MW. Ten of the reactors will be 500-MW units (now in the design stage) and the rest will be similar to the 235-MW plants in operation in Madras and Rajasthan. The 15-year plan would cost $14,000 million at 1983 prices. Because government resources are limited, it has been decided to bring nuclear power generation under a public limited company that will raise funds through the issue of bonds and shares. Although AEC is confident of launching the Nuclear Power Corporation, it remains to be seen whether the Indian public will invest in this programme after Chernobyl.

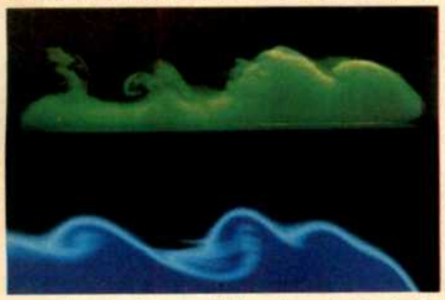
K.S. Jayaraman

Storm flow simulation by computer

Washington

A numerical simulation of thunderstorm outflows by Kelvin Droegemeier of the University of Oklahoma is being used to program a 727-passenger jet flight simulator at the Federal Aviation Administration's training academy in Oklahoma City. It is hoped that pilots can learn to recognize and avoid dangerous wind shear associated with outflows, which has caused several fatal crashes.

Droegemeier has produced a film of his simulations, which were run on a Cray X-MP supercomputer at Omnibus Simulation Inc. of Los Angeles. The blue image shows a still from the numerical model; the green image shows a laboratory tank study in which the same phenomenon is modelled by releasing salt water into less dense fresh water. The simulation, which has about 40,000 grid points, reproduces clearly the features seen in tank studies 20 years ago. The turbulent eddies on top of the outflow in both images are corroborated by Doppler radar images of real outflows, and probably explain the seemingly random pressure fluctuations observed on the ground beneath. Production of Droegemeier's film was paid for by the National Science Foundation's Office of Advance Scientific Computing.



Tim Beardsley

AIDS

German survey's gloomy outlook

Three-quarters of those infected with the AIDS (acquired immune deficiency syndrome) virus will enter the final and fatal stages of the disease within seven years, according to a research team from the University of Frankfurt.

The predictions from the study, claimed to be the first long-term project in Europe, are based on "continuous observation" of 543 patients, largely the sexual partners of the first AIDS victims to die in Frankfurt, who have been attending outpatients' clinics since 1982. On the basis of the rate of deterioration of these patients — 377 proved positive in original antibody tests — and from computer forecasts, the researchers conclude that half the antibody carriers will have progressed to the final stage of AIDS within five years of first exposure and three-quarters within seven years.

The research team is led by Professor Eilke Brigitte Helm from the Zentrum der Inneren Medizin at the University of Frankfurt and the findings are published in the *Deutsche Medizinische Wochenschrift* (No. 32/32, 1986). The team placed all the patients in one of four categories, depending on the seriousness of the symptoms. They were those whose antibody tests proved positive but who had no obvious sign of the disease; those with a slight immunological weakness and showing symptoms of lymph-node swelling; those with pronounced immunological defects; and those with the full symptoms of the disease. The progress of the patients was monitored for three years.

At the end of that period, only 30 of the original antibody-positive patients were still healthy and, within a year, half of the patients with no symptoms moved into the next category. Within two years, half of those with slight or pronounced immunological defects were entering the final stages of the disease or suffering severe disintegration of their defence systems. With one exception, all those in the final stages of the disease died within two years.

According to the researchers, the belief that AIDS would develop only in a minority of cases infected with the virus is unfounded. The predicted size of that minority has increased with time from 5 or 10 per cent to at least 30 per cent in some cohorts of patients, but the German prediction of 75 per cent is unprecedented.

Bill Johnstone

Technology

Genentech buys out partnerships

Washington

GENENTECH Inc., one of the leaders in US biotechnology, has decided to reap an early harvest by announcing plans to buy out two of its research and development partnerships early. A limited partnership is a means by which outside investors support part of a company's research and development programme, sharing the profits if it succeeds and otherwise being able to set off the loss against taxes.

Genentech Clinical Partners Ltd (GCP) and Genentech Clinical Partners II (GCP II), have provided support and shouldered risk while Genentech's seedling products, including the human growth hormone Protropin, tissue plasminogen activator Activase and gamma interferon are being developed and tested clinically.

According to the original agreements, Genentech would not have bought out the partnerships until two years after the products from each of them had received approval from the US Food and Drug Administration (FDA). This would have been 1988 for GCP, as FDA approved Protropin this year, and a later date for GCP II, depending on FDA approval for the marketing of Activase.

Buying out the partnerships will give Genentech exclusive US rights for Protropin, Activase and gamma interferon, and all the associated profits. The partners will be offered 3,000 shares of Genentech common stock for every $50,000 partnership unit in GCP and 2,500 shares for every similar unit in GCP II. If the stock price on the date of the transactions remains at $80, as expected, each $50,000 investment in GCP would receive Genentech stock valued at $240,000 and in GCPII stock valued at $200,000.

By acting now, before the partnership agreements have expired, Genentech avoids buying completely developed products that would have to be capitalized and amortized. This development shows Genentech's faith in the products developed under the partnership agreements; on Wall Street, Genentech's stock rose by $12.75 per share on the day after the proposed buyout was announced. The price is now about $85.

Genentech is one of the first US biotechnology companies to become profitable. It recorded a profit of $5.6 million ($0.18 per share) in 1985, and its financial report to the end of September this year shows a net income of $8.2 million, an earning for each share of $0.23. Although the early promise of big profits for 1986 will be diluted by the one-off expense of the GCP II buyout, profits for 1986 are likely to be double those for 1985.

Genentech is also now planning a two-for-one stock split and intends, like many other leading US companies, to reincorporate in the state of Delaware to benefit from the favourable legal climate there.

Carol Ezzell

Student squeeze in West Germany

Hamburg

WHILE the universities in West Germany are still groaning under the burden of too many students — 1.3 million are sharing 850,000 places (see *Nature* 322, 197; 1986) — they are preparing to compete for students in the decade ahead. The number of school students is falling; the era of expansion is coming to an end.

A detailed analysis carried out at the centre for the study of regional development at Giessen and sponsored by the Deutsche Forschungsgemeinschaft (DFG) has shown that the new student generation will prefer big and old universities, so that the new and smaller institutions, particularly in Bavaria, may "have to fight for their survival".

The study is based on official university statistics, interviews with school leavers and figures from the Zentralstelle für die Vergabe von Studienplätzen (ZVS). The results are alarming for the cities concerned. Giessen, for example, with 17,500 students among its 70,000 inhabitants, will lose more than 4,000 students.

Other universities, many of which are at present overcrowded, will not suffer so greatly. The analysis, which was carried out by Professor Ernst Giese, concluded that 22 out of 52 West German universities are "above average", in being chosen by more than 2 per cent of new undergraduates. The University of Munich attracts 8 per cent, followed by Cologne, Hamburg, Hannover, Münster, Munich (Technical University), Aachen, Bonn and Berlin (Free University) with between 5.49 and 3.5 per cent.

Old universities are preferred because of their traditions but also because of the recommendations of previous students. Universities such as Bremen, Frankfurt and Marburg, which were considered to have left-wing tendencies are well regarded now.

One of the ways in which the universities will compete is by setting up special study centres concentrating on specific subjects such as solid-state physics at Stuttgart, genetic engineering at Heidelberg and Würzburg, plant biotechnology at Hamburg and Cologne and information technology at Karslruhe. Jürgen Neffe

Chinese – US collaboration

NSF refuses fees

Washington

THE National Science Foundation (NSF) is digging in its heels against what it sees as unreasonable charges levied by the Peoples' Republic of China for US field trips in the country. Officials say the Chinese policy of charging for field trips is contrary to the spirit, if not the letter, of government agreements on scientfic exchanges, and have given notice that NSF will not pay for a return visit next year by a US team taking glaciological samples in the Quilian mountains in Tibet.

The US team, headed by Lonnie Thompson of Ohio State University, spent 28 days last summer in a remote glacial area 1,000 km west of Lanchow, working with scientists from the Lanchow Institute of Glaciology and Geocryology. According to NSF's Eugene Bierly, China at first asked for $100,000 to defray expedition expenses, but after negotiations, the figure was reduced to $80,000. Thompson plans to return next summer to bring back ice-core samples, but NSF says it will not meet field-trip expenses again on the grounds that they violate the "receiving side pays" convention that underlies the US – Chinese agreements.

Bierly said there would be no question of charging Chinese scientists for field trips in the United States, and expressed the fear that budget cuts at some Chinese research institutes might tempt them to make deals with scientists from other countries, perhaps allowing priority of access to those who pay well. Bierly said NSF is determined to "hold the line".

Others report similar difficulties with fees for access to China. Craig Black of the Los Angeles Museum of Natural History said last week at a meeting of the National Science Board that charges for biological field trips were commonplace and that it was "almost becoming competitive". Officials at the meeting were alarmed by reports of a Japanese–Chinese agreement on glaciological research according to which, it is said, Japan will pay $1,000,000 and donate four expedition vehicles in return for access to a large tract of China. NSF officials say the foundation would be unable to offer funds on that scale.

Thompson disputes NSF's figure of $80,000 for the fee for last year's Quilian expedition, saying it was only $40,000. The fee was, he says, reasonable for the level of logistical support provided, and that it "would be a good deal in any other country". NSF meets expedition costs in countries with which there are not collaborative agrements. But Thompson says he has agreed with the National Geographic Society that it will pay the expedition costs next year, so NSF's stand should not imperil his venture. **Tim Beardsley**

Rhine pollution

Death of Europe's sewer?

In the French city of Strasbourg, on a tributary of the Rhine just 100 km downstream from Basel, feelings are running high, as they are all along this long river. Demonstrators gathered in Strasbourg and other riverside towns on Sunday carrying placards with messages such as "Poison ou poisson: il faut choisir" ("Poison or fish: we must choose"). Some fish have survived, but are polluted, and fishing has been banned for six months along the upper Rhine and its canals.

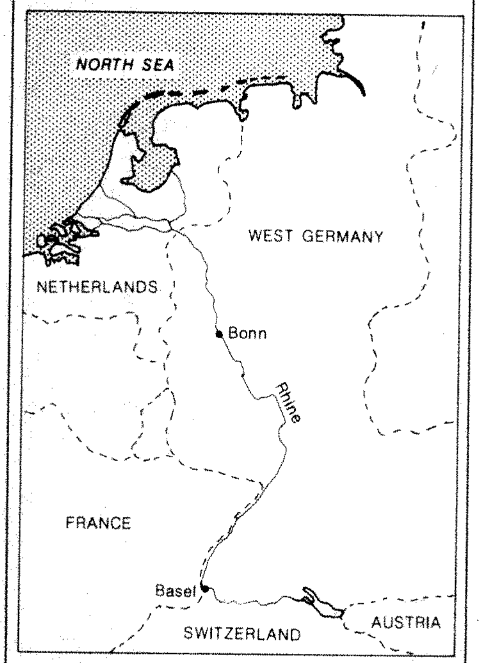
The Rhine seems 'dead' for 100–200 km downstream of Basel in northern Switzerland after the fire in the Sandoz chemical plant led to the discharge into the river of up to 32 pesticides, fungicides and other agricultural chemicals. One of the most important, 'disulfeton', is said to be twice as toxic as potassium cyanide; the average dose to kill 50 per cent of rats (the LD_{50} criterion) is about 5 mg kg^{-1} compared with 10 mg kg^{-1} for potassium cyanide.

These facts emerged after the owners of the factory made public a list of the 1,000 tonnes of chemicals stored on the Basel site. Disulfeton has been detected as the principal pollutant from the accident as far as the Dutch border, 600 km downstream from Basel; but by this point the disulfeton had been greatly diluted to around 2 mg l^{-1} of Rhine water, according to informed sources.

Sandoz and the Swiss Federation have admitted liability for the accident, which occurred after firemen had flooded Sandoz's Basel factory with water during a fire, to the extent that a dyke was breached and a toxic 'soup' of some 10–30 tonnes of agrochemicals was released into the Rhine. Mercury-based fungicides such as tillex as well as disulfeton and other organophosphorus compounds were in the mixture, which killed fish, normally resistant eels and insects. The effects on plant life are as yet unknown, and although little organo-mercury was detected downstream, there are fears that such chemicals have been deposited as sediment close to the release point, where they may be slowly released to the food chain for some time.

Thus, according to European environmentalists, the damage has been "enormous". According to the European Environmental Bureau, an organization that acts in Brussels for a number of European environmental groups, the Rhine was beginning to recover until this accident. Industry had been reducing its emissions under a combination of voluntary and statutory controls, to the point where fishermen were beginning to look for salmon at Basel. But this accident released as much pollution in a couple of hours as the Rhine normally receives in a year. "It's set the clean-up back a decade", a bureau spokesman said.

Meanwhile, people are despondent about the months ahead. The proprietor of one fishing tackle shop said she had had no customers for 3–4 days and expected none now until the spring. Did she expect any compensation from the authorities? She gave a resigned shrug of the shoulders.



Nevertheless, there is no doubt that the accident has set the European environmental movement and the legislators a good few steps forward, just as the Chernobyl accident gave a boost to the anti-nuclear lobby. Last week, the European Commission was pointing out how many countries have failed to put into effect the 'Seveso directive' of 1982, in which European Community countries agreed (subject to the passage of relevant national legislation) to impose controls on chemical factories. This followed an accident at Seveso in Italy in which a large area was contaminated by dioxin.

But four years after the directive, the Commission is satisfied only with legislation enacted by Denmark, France and Britain. The Netherlands, one of the loudest complainants against the Sandoz accident, has failed to apply the Seveso directive's principles on worker protection; West Germany has problems in the application of rules on the storage of dangerous substances; Greece's legislation has been too vague and "inadequate"; and Italy, most affected by Seveso, has not applied one of the directive's most crucial elements, a requirement to be placed on companies regularly to inform the authorities and their workers of the processes taking place in a factory, of the chemicals stored and of emergency plans.

The Commission is now planning action against defaulting countries, but to begin with is only taking on a minnow — Luxembourg — in the European Court; with the rest, it has satisfied itself with sending stern diplomatic notes.

Meanwhile, countries belonging to the governmental International Commission for the Protection of the Rhine met in Zurich last week, heard contrite Swiss apologies and offers of damages (to a scale as yet undetermined), and set in motion the investigation of measures that would considerably tighten up controls and safeguards on Rhine industries. Switzerland itself will consider adopting national legislation along the lines of the Seveso directive of the European Communities — although as a non-member it has no obligation to do so — and there is now great pressure for Switzerland to sign an international convention that would oblige Swiss companies to inform their national neighbours of their activities, safety procedures and contingency plans.

The members of the International Commission have also decided in principle to apply to the Rhine the principles of the European convention on the reporting of oil releases at sea (Switzerland was very late in informing its downstream neighbours of the Sandoz release) and to set up an insurance system equivalent to TOVALOP, the tanker owners' voluntary agreement on liability for oil pollution. Any new conventions of this kind would of course apply to all forms of pollutant release and not just to oil. **Robert Walgate**

• Chemical accidents and spills are not confined to Western Europe. During the past few weeks, several incidents have been reported from the Comecon bloc.

On 1 November, at the chlorine and PVC works in Devnya (Bulgaria), there was an accident in which several people were killed, which destroyed the control rooms of two sections of the plant and also the finished products store, and which was apparently caused by rupture by an inadequately maintained pipeline.

On 11 November, in the Buna chemical works in East Germany, a hot-condensate condenser was destroyed by excess pressure, with four workers seriously injured and "considerable damage to property". And on 12 November, Prague radio reported that an accident in the Ostrava sewerage system had led to the discharge of heating oil into the rivers Lucina and Ostravica, and then into the Oder. The damage, the Chechs said, was expected to be long-term because of contamination of the river banks, but "drinking water has not been affected". According to the East German Ministry for Environmental Protection and Water Management, the spillage was not expected to affect the East German reaches of the Oder. **Vera Rich**

Genetic manipulation

Living outside regulation

Washington

A GENETICALLY altered strain of yeast designed to produce a low-calorie beer will soon be churning out trial brews at the Brewery Research Foundation (BRF) in Surrey, England. BioTechnica International of Cambridge, Massachusetts, chose to test its new creation outside the United States in part to postpone facing the regulatory process awaiting it in this country.

BioTechnica has so far produced only 10 millilitres of beer in the laboratory to test the process, and is prohibited from producing more than 10 litres by the Cambridge Biohazards Committee. To do testing on a larger scale, BioTechnica sought out BRF in England because there is no equivalent brewing test facility in the United States. But Roger Yocum, director of BioTechnica's yeast research division, makes no bones about the fact that testing the new yeast will probably be much easier in England.

BioTechnica International has been working for more than three years to produce a strain of *Saccharomyces uvarum* (also called *S. carlsbergensis* or brewers' yeast) containing a gene from the mould *Aspergillus niger* encoding for a glucoamylase enzyme. So-called lite (for 'light') beers are less fattening because they contain less starch than standard brews. By introducing a glucoamylase into brewing yeast, the beer can be produced faster and without other additives to remove the starch.

John Hammond, a biochemist at BRF, says that so far all work on the yeast has involved less than 5 litres, and all brews have been pasteurized to kill the yeast. But an unpasteurized beer, such as is provided to bars and restaurants, could contain live yeast, raising the issue of environmental release of what is an intergeneric organism. Hammond says BRF will consult with the Advisory Committee on Genetic Manipulation and other relevant governmental agencies before proceeding with any large-scale work with BioTechnica's yeast. But Yocum hopes that any US decision whether to permit environmental release of the new yeast will be clarified by BRF's testing.

David Kingsbury, assistant director of the National Science Foundation and one of the designers of the new government regulatory framework for biotechnology, says that even though the new strain of *S. uvarum* is probably safe and not an environmental hazard, it would come under the highest level of federal scrutiny because it is an intergeneric organism. Kingsbury says this shows that the new federal guidelines are if anything too strict, rather than too lax as some critics have charged. In the case of the new yeast, the Food and Drug Administration (FDA) would probably have initial regulatory authority. Henry Miller, special assistant to the commissioner of the Food and Drug Administration for biotechnology, says the fact that the yeast is a intergeneric organism would have no particular relevance to FDA's regulatory process. FDA's charge is to determine whether it is a food additive, and whether it transforms a product that is already "generally recognized as safe".

Testing outside the United States is hardly a panacea, as the Wistar Institute in Philadelphia, Pennsylvania learned recently. Wistar and the Pan American Health Organization (PAHO) had been conducting a field test of a rabies vaccine using a modified vaccinia virus containing the sequence for an immunogenic protein from the rabies virus. Twenty cows were injected with the modified virus — ten subcutaneously and ten intramuscularly — and housed with twenty unvaccinated animals.

In September, news of the experiment appeared in the Argentine press, stirring a political controversy. That PAHO had not informed the Argentine government that the field trial involved a recombinant DNA organism made matters worse. PAHO has been conducting trials on rabies vaccines at a field station in Azul, 180 miles south of Buenos Aires, and it is unclear whether they had a legal obligation to inform Argentine officials.

Privately, a PAHO official concedes it would have been politic, given the politically sensitive nature of recombinant research, to inform Argentine authorities. Rabies is a significant problem in Argentina, with millions of dollars is cattle lost annually from the disease. Preliminary results from the Azul experiment show that the inoculated animals did show an immune response to rabies, while the uninoculated animals and the four human caretakers did not.

The furore in Argentina is ironic as Wistar associate director Warren Cheston says an Argentine government research laboratory contacted the Wistar Institute in 1984 to ask about conducting field trials with the vaccine. At that time Wistar declined, saying additional laboratory testing was needed.

A team from the Oregon State University in Corvallis had better success testing its modified vaccinia virus. The university announced last week that a field trial for a vaccine for Sindbis virus had been carried out over the past year in New Zealand. Calves, sheep and chickens were vaccinated in the study. None developed disease or transmitted the virus to unvaccinated animals. **Joseph Palca**

India

Move to stop sex-test abortion

Bangalore

CONCERNED over the steady increase in the incidence of female feticide, the government of India is planning legislation to restrict the use of techniques for determining the sex of a fetus. The central health and family welfare ministry in New Delhi is seeking advice from medical experts, social workers, lawyers and women's groups in formulating the legislation. The aim is a comprehensive bill that would drastically limit the practice of amniocentesis and other emerging techniques of prenatal sex-detection.

As it is, the Medical Termination of Pregnancy (MTP) Act of 1971, which stipulates that an abortion can be performed only if there is risk to the life of the pregnant woman or can cause grave injury to her mental or physical well-being, has been ignored by some of the medical fraternity.

For many months, women's groups in Bombay and New Delhi have not only been campaigning against sex-linked abortion but have also been demanding that the authorities provide effective legislation to stop the abuse and to prevent the spread of clinics that exploit the legal loophole.

Mr R.P. Ravindra, a social worker in Bombay, points out that amniocentesis as a tool for sex-determination is almost always followed by abortion if the fetus is female. He says that a study of 8,000 cases of abortion showed that 7,997 of the fetuses were female.

Altering the sex ratio (at present 935 females to 1,000 males) through the abortion of female fetuses could adversely affect the social balance. India has one of the highest maternal mortality rates in the world. Moreover, about 70 per cent of women in India suffer from anaemia and malnutrition, both of which are likely to be aggravated by repeated abortions after sex-determination.

Supporters of the tests claim that the number of unwanted and neglected female children will be reduced and that a scarcity of women will raise the status of women in general. Others believe that the result will be polyandry and an increased incidence of rape and prostitution.

Not much is known about the numbers of sex-test clinics or their operation. Many are run under the guise of maternity homes, clinical laboratories and family health centres. Many are in the industrially and agriculturally advanced states of Maharastra, Punjab, Haryana and Delhi. There are 248 clinics in Maharastra according to the state health department. **Radhakrishna Rao**

Japan

Body light points to health

Tokyo

THE Research and Development Council (RDC) of the Japanese Science and Technology Agency has sufficient confidence in Professor Humio Inaba of Tohoku University to invest about Y15,000 (£70 million) in his research. For Inaba believes that the light people emit may reflect the state of their health.

Inaba's team has found that the blood of patients suffering from diseases such as cancer, diabetes and jaundice gives off much more light as that of healthy people. Smokers' blood "glows" twice as brightly as that of non-smokers, but returns to normal after a day of not smoking.

Although the mechanism of photoemission in each of these cases has still to be determined, peroxidation of lipids through the generation of activated oxygen and free radicals is believed to be involved. Peroxidal lipid reactions are suspected agents in diabetes, liver and lung diseases, arteriosclerosis, ageing and cancer. Inaba believes biophotons will eventually be used for diagnostic and clinical purposes.

The biophoton project is one of the latest additions to Japan's ERATO programme, which was set up in 1981 in a radical attempt to break free from Japan's conservative research system.

ERATO projects run for five years and, with the sole exception of the project leader, the 20 or so participants are all young researchers (under 35) drawn equally from universities, government research institutes and industry. Two or three foreigners are usually included in the team and the research is carried out in independent laboratories uncontaminated by the usual environment of Japan's academic and government research laboratories. Once the theme of the project is chosen, the teams are left to find their own path, sometimes with bizarre consequences.

Inaba, who is attached to Tohoku University's Research Institute of Electrical Communication, became involved in maser research in the late 1950s, during an unsuccessful bid to get a large radiotelescope built in the Tohoku area north of Tokyo. In the 1960s, he turned to the development of high-sensitivity detectors for lasers, and by the end of the decade had devised a laser system for the detection of air pollution using highly sensitive photomultipliers to pick up the scattered light (*Nature* **224**, 57; 1969).

Using Inaba's equipment, biochemists trying to detect the low levels of light expected to be given off during biochemical reactions involving peroxidation were able to detect ultraweak light emissions (of the order of tens of photons per second) from tumours and regenerating livers in mice and rats. But dark noise prevented determination of the spectrum of the emitted light.

Inaba's team, with support from RDC, went on to develop a luminescence analyser in collaboration with Tohoku Electronic Industrial Co. Ltd that can detect less than one photon per second and can handle samples up to the size of a rabbit.

The project will aim at improving this technology to allow measurement of weak biophoton emissions and to bring spatial resolution ot the sub-micron level.

The technology has already found applications in the salad oil, rubber, paint and powdered milk industries, where the degree of oxidative deterioration of the product can be checked in a few minutes by monitoring ultraweak chemiluminescence. But medical applications seem to hold most promise.

David Swinbanks

SERC annual report

THE UK Science and Engineering Research Council (SERC) has asked the government, through the Advisory Board for the Research Councils, for an extra £20 million to make up the shortfall in its budget for 1987–88 due to unfavourable fluctuations of the pound against European currencies. The disclosure was made last week on the publication of SERC's annual report.

Professor E. W. J. Mitchell, chairman of the council, in the introduction to the report, claimed: "Overall the council's task remains the achievement of a sensible balance between curiosity-driven basic research, the seed corn of the future, and the strategic work where application might be anticipated in the medium term. This task is made more difficult within a severely contrained budget, further exacerbated by the wayward behaviour of the exchange rates which in 1985–86 were invariably against us."

Bill Johnstone

Nuclear safety

NRC investigates Seabrook

Boston

THE alleged misuse of drugs and alcohol by construction workers may have compromised the safety of the Seabrook Nuclear Power Plant in Seabrook, New Hampshire, which is due to begin low-power testing at the end of this year. The Nuclear Regulatory Commission (NRC) is now investigating these allegations.

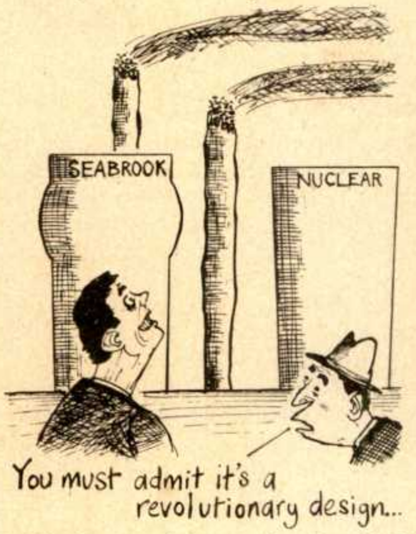
The outcome of the inquiry is likely to affect both the debate about mandatory drug testing and that about existing controls on the safety of nuclear power plants. Drug tests are opposed by those who consider them an infringement of personal liberty.

Several groups, including the energy conservation and power subcommittee of the House of Representatives headed by Congressman Edward J. Markey (Democrat, Massachusetts), have gathered allegations from police, hospital personnel and court employees as well as from the construction workers themselves. The alleged abuses are said to have transformed the town during the ten years of construction.

Former plant workers called Seabrook "a drug addicts' haven". Police have reported finding steel workers "drunk, coming and going" and, according to Dr Gary Lamphere of Exeter Hospital, "it was unusual enough in the early 1980s to see a Seabrook patient without alcohol on his breath to remark on it".

The issue is not whether workers constructing nonvital areas of the nuclear power plant were drunk or high on drugs, says Thomas Elsasser of NRC's Pennsylvania office, but whether "systems essential to safety" have been affected. NRC will judge Seabrook's safety by inspecting documents, examining systems on a sampling basis and monitoring containment and fuel-load tests.

So far, Seabrook has passed all tests: 21,000 hours of inspection by NRC, almost twice that for any other plant in the region, have uncovered no structural reason to believe the plant is unsafe.



You must admit it's a revolutionary design...

But testing after construction is considered unsatisfactory because some parts of the plant are no longer accessible for inspection. NRC may face other obstacles: ti is alleged that some quality control inspections were not completed properly nor documented correctly.

Elizabeth Collins

Queensland and Creationism

SIR—In response to inaccurate statements made by Tony Thulborn (*Nature* **315**, 89; 1985), I am writing to clarify my department's policy on the teaching of the theory of evolution in Queensland State secondary schools.

The Queensland Education Department policy on the teaching of the theory of evolution requires that it be taught (1) as a theory, not as fact; (2) in a balanced manner.

In providing balance, teachers must acknowledge that alternative theories to evolution exist. Some may be supported by scientific evidence; others may not be scientific in nature, but nevertheless based on beliefs deeply held by a significant proportion of the community.

Beyond this, I believe that the question of what constitutes "balance" should be left to the professional judgement of each and every science teacher. I hasten to add that I would be most concerned if a teacher taught the theory of evolution in such a way as to assail the religious beliefs of students. I would be equally concerned if the intent of a science syllabus, in terms of both "content" and "minimum time", were not adhered to.

LIN POWELL
(Minister for Education)
Treasury Building,
Queen Street, Brisbane, 4000,
Australia

Sexism and conservation

SIR—The advertisement on the back cover of the 16 October issue of *Nature* shows a young woman clad in only a shell necklace. While past correspondence has discussed the issue of sexism in advertisements in your journal, we want to point out that many gastropods are fished and killed for the explicit purpose of making jewellery because local entrepreneurs have exhausted the dead shells that are washed up on beaches. In certain countries, especially those with tropical shorelines, this practice has reached epidemic proportions and constitutes a major threat to the resident molluscs[1]. Display of material that links (live) girls to (dead) prosobranchs can only damage efforts to conserve sea-shore fauna.

T. M. CARE
Sub-Department of Animal Behaviour,
University of Cambridge,
Madingley, Cambridge CB3 8AA, UK

W. K. LINDSAY
Department of Applied Biology,
University of Cambridge,
Pembroke Street,
Cambridge CB2 3DX, UK

1. Kendall, B. *Swara Magazine* **8** (1), 8–11 (1986).

Literary pedantry

SIR—In "What is the scientific literature" (*Nature* **322**, 681; 1986), John Maddox appears to be confusing the medium with the message. Although I, too, decry the poor writing that characterizes many scientific publications, this alone does not drum papers out of the ranks of 'literature'. On the contrary, any scientific paper meets all standard dictionary definitions of 'literature', with the possible exception of the definition that is characterized as *Archaic*, namely, for example, "literary culture" (*Dictionary of the English Language*. Random House, New York, 1968). Maddox is barking up an antique tree.

MELVIN BLECHER
Department of Biochemistry,
Georgetown University,
Washington, DC 20007, USA

Problems of Soviet refusniks

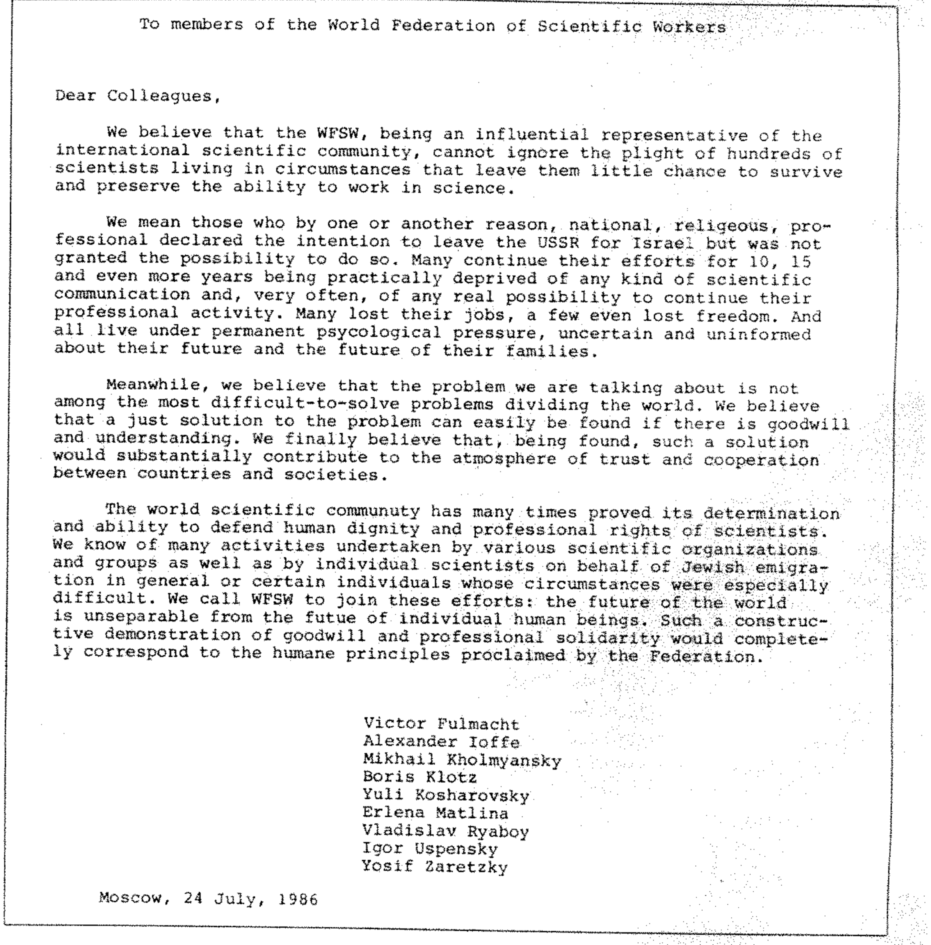
SIR—We beg the courtesy of your columns to address a letter, signed by us and seven of our colleagues, to members of the World Federation of Scientific Workers (WFSW) and its president, Professor Jean-Marie Legay.

We had hoped to hand this letter personally to Professor Legay when he was in Moscow this summer for the General Assembly of WFSW. Indeed, when we telephoned him there he courteously agreed to meet the two of us at his hotel on the morning of 24 July, before the opening ceremony of the assembly. On arrival at the hotel, however, we were denied entry and had to telephone Professor Legay again to explain our difficulty, which he promised to try to deal with. Instead, after some delay, we were forced out of the hotel and away from the entrance area by a group of plain-clothes men who brutally informed us that we had "no reason to meet Legay". Subsequent telephone conversations established Professor Legay's inability to arrange a meeting with us during the remainder of his stay in Moscow.

The text of the letter appears below.

We do not wish to embarrass Professor Legay and WFSW. Indeed, what happened on 24 July was as great a discourtesy to them, in being denied a visit by their invited guests, as it was an affront to us. Instead, what we wish to do is give WFSW, committed as it is to professional solidarity among scientists, the opportunity publicly to press for our professional rights and thus enhance its prestige as an organization dedicated to humane principles.

ALEXANDER IOFFE
BORIS KLOTZ
Moscow, USSR



To members of the World Federation of Scientific Workers

Dear Colleagues,

We believe that the WFSW, being an influential representative of the international scientific community, cannot ignore the plight of hundreds of scientists living in circumstances that leave them little chance to survive and preserve the ability to work in science.

We mean those who by one or another reason, national, religeous, professional declared the intention to leave the USSR for Israel but was not granted the possibility to do so. Many continue their efforts for 10, 15 and even more years being practically deprived of any kind of scientific communication and, very often, of any real possibility to continue their professional activity. Many lost their jobs, a few even lost freedom. And all live under permanent psycological pressure, uncertain and uninformed about their future and the future of their families.

Meanwhile, we believe that the problem we are talking about is not among the most difficult-to-solve problems dividing the world. We believe that a just solution to the problem can easily be found if there is goodwill and understanding. We finally believe that, being found, such a solution would substantially contribute to the atmosphere of trust and cooperation between countries and societies.

The world scientific communuty has many times proved its determination and ability to defend human dignity and professional rights of scientists. We know of many activities undertaken by various scientific organizations and groups as well as by individual scientists on behalf of Jewish emigration in general or certain individuals whose circumstances were especially difficult. We call WFSW to join these efforts: the future of the world is unseparable from the futue of individual human beings. Such a constructive demonstration of goodwill and professional solidarity would completely correspond to the humane principles proclaimed by the Federation.

Victor Fulmacht
Alexander Ioffe
Mikhail Kholmyansky
Boris Klotz
Yuli Kosharovsky
Erlena Matlina
Vladislav Ryaboy
Igor Uspensky
Yosif Zaretzky

Moscow, 24 July, 1986

Mathematics for natural hierarchy

A new way of tackling the statistical problems of spin glasses relies on almost forgotten mathematics, but may have a bearing on problems in biology still unsolved.

How does it so often arise that the solution of newly defined problems hangs on the use of mathematical techniques independently developed, often decades earlier, for quite different purposes or (mathematicians being what they are) for no purpose at all? That question is the starting point for a delightful article in *Reviews of Modern Physics* for July this year in which the concept of "ultrametricity" is offered as an illustration of the principle that essential tools will usually be found lying discarded in the attics of the mathematicians. It goes without saying that the real world is not always as bountiful.

Part of the charm of this winsome piece is that its authors, R. Rammal (from the CNRS low-temperature institute at Grenoble), G. Toulouse (from the École Supérieure de Physique et Chemie Industrielles in Paris) and M.A. Virasoro (of the physics department of the University of Rome), confess that they themselves have only recently stumbled on the notion of ultrametricity, in an article on spin glasses in another physics journal. Like new converts in most fields, they have become enthusiastic proselytizers for the cause (*Rev. Mod. Phys.* **58**, 765; 1986).

Historically, so far as Rammal *et al.* have been able to tell, ultrametricity goes back to 1897, when the German mathematician Kurt Hensel published a recondite piece of algebraic number theory showing, among other things, that an arbitrary integer can always be represented as a sum of ascending powers of some other arbitrary integer, with each term in the series multiplied by a coefficient which may be zero but which is, in any case, less than the base for the ascending power series.

Nowadays, the most familiar illustration of the truth of Hensel's principle is the practice well known to computer buffs of representing decimal integers in binary form, which is merely a way of representing an arbitrary integer as a sum of powers of 2 in which sum the coefficients may be either 0 or 1. The more general statement is that the base of daily arithmetic may be changed from 10 to any other arbitrary base. A further generalization is that rational numbers (with finite, or non-recurring, decimal representation) may always be represented by the sum of powers of some arbitrarily chosen base in which some negative as well as positive powers are included. Rammal *et al.* say that Hensel's work was done in complete isolation, but its practical importance now cannot be gainsaid.

The other mathematical innovation from which ultrametricity has sprung is the formal definition of metric spaces at the turn of the century. For a 'space', or a set of points, to be metric, there must be some function called distance of any pair of points which satisfies the condition that the value of the function is zero when the two points are the same, and that the value is unaltered when any pair of points is interchanged.

Euclidean verities come to mind, together with the euclidean principle that the sum of the distances between one point and a second and between the second and third is greater than the direct distance between the first and the third. This is just another way of saying that, in the euclidean world, the sum of the lengths of two sides of a triangle (say AB plus BC) is greater than the length of the third side (AC). Ultrametricity requires a more stringent principle, that AC should be less than whichever is the greater of AB and BC.

Such a metric world is, of course, an odd world. Rammal *et al.* point to some formal mathematics in the 1940s showing that, with definitions of metric distance satisfying these conditions, any point inside a sphere (defined as the region distant less than some fixed amount from a fixed point) will be itself at the centre of the sphere, and that the only possible triangles will be either equilateral or isosceles (in which case the base must be shorter than the two equal sides). Hensel's work is relevant because it provided one of the first illustrations of how such a metric might be defined.

What relevance can such an odd way of defining distance have to the problems of the real world? Rammal *et al.* start with a phylogenetic tree such as any cladist might construct. The tree implies a hierarchy which, once drawn, implies metric distances between the members at the lowest level, most easily thought of as the length of time since any pair of them diverged from a common ancestor. Then, in the comparison of the distances between three arbitrary extant species, either one pair has more recently evolved from a common ancestor than that evolved from some more distant common ancestor (in which case the triangle is isosceles) or no two of the three species have a common ancestor earlier than the putative common ancestor (in which case the triangle of the distances between them is equilateral in the ultrametric sense).

From this point, it is an easy step to the application of the ultrametric concept to all kinds of intriguing problems in data processing, especially to the use of tree-sorting algorithms. But Rammal and his colleagues are proudest of their novel application of ultrametricity to the problem of spin glasses, the structures (real enough) represented by three-dimensional lattices whose vertices are occupied by spins (think of them as little magnets) and whose energy is the sum of the magnetic interactions between all possible pairs of vertices in the lattice. The problem is that of an Ising lattice in which the simplifying assumption that only nearest neighbours interact is abandoned.

Experimentally, the spin glasses are peculiar materials, especially at low temperatures. Phase transitions occur to states in which the solid material appears to consist of small domains of magnetically ordered spins, but the degree of hysteresis associated with changing external circumstances (such as temperature) is unusual; the materials seem to have a capacity for long-term memory of their previous state. In the decade since the problem of the spin glasses was recognized as such, much has been done by analytical and numerical approximations to show that the low-temperature condition of such a material may be one of a great number of essentially similar configurations that are distinguishable from each other largely by their symmetry. What Rammal *et al.* are able to show is that these configurations are related to each other hierarchically, by means of an ultrametric metric.

That will be an interesting landmark, but will it be an important one? Much depends on whether the authors' speculations turn out to be realistic. They claim, for example, that their line of argument may have a bearing on the travelling salesman problem (how to travel to N cities most economically), the founding conundrum of linear programming. Naturally, they say there is a place for neurobiology in their scheme of things. And since the folding of protein molecules into biologically active materials is a hierarchical process, why not subsume that as well? If they are right, a great many people will be learning a great deal of mathematics they did not previously know existed.

John Maddox

Biological conservation

National park boundaries and ecological realities

David S. Wilcove and Robert M. May

WHEN the United States Congress established in law the boundaries of Yellowstone National Park in 1872, the main motivation was the geology of the area, not the biology. The legislators wanted to create a park of sufficient size to protect the remarkable geothermal features[1] and the boundaries were drawn more or less as a giant rectangle, measuring about 99 by 86 kilometres, and containing well over 900,000 hectares of land. Almost immediately, some people began to wonder whether 900,000 hectares was sufficient to preserve the park's wildlife and natural features. Today, its population of grizzly bears is declining; each winter, elk move out of the park to a much reduced winter range; the state of Montana, fearing the spread of brucellosis to domestic livestock, wants to eliminate any bison that wander off the park; and geothermal energy development outside the park boundaries threatens to disrupt its geysers. A recent study[2] suggests at least seven out of eight national parks in western North America are too small to maintain minimum viable populations of several species found within them.

The lack of congruence between legal and ecological boundaries[2] is common to many parks. In Kenya, the seasonal movements of elephants and other game species take them well outside the boundaries of Amboseli National Park (39,200 hectares). When outside the park, the animals must compete with livestock and contend with poachers. In Costa Rica, three-wattle bellbirds (among other species) annually migrate from the protected montane forest of the Monteverde Reserve into unprotected lowlands. Where park boundaries cross ecological realities, it is usually because the park is too small relative to the space required by particular species, or too small relative to the scale of the natural and man-made disturbances that occur in and around it, such as fires, disease epidemics, insect outbreaks and seasonal or other variations in food supplies.

Territories

Large mammals, like the Yellowstone elk or the Amboseli elephants, need a lot of food and water and must travel far to get it. Their seasonal movements often encompass an area far greater than that of the parks in which they occur and illustrate the particular difficulties posed in the conservation of migratory species. Large predators, which usually occur at low densities, present different problems. As some grizzly bears have home ranges of over 3,000 km^2, Yellowstone National Park by itself simply could not support a viable population of them. The park's grizzly bears are just a part of the larger population that includes those bears in the adjacent national forests.

In an attempt to assess this problem, Newmark[2] recently compared the legal and ecological boundaries of eight national parks in western North America. He defined the ecological ('biotic' in his terminology) boundaries as the area encompassing the entire watershed of a park enlarged, if necessary, to encompass an area of sufficient size to maintain a 'minimum viable population' for the terrestrial non-flying species with the largest home range found within the legal boundaries. Setting the minimum viable population size at 500 — which is probably an underestimate — Newmark found that seven of the eight parks were too small by factors ranging from 6 to 100. Even when the level was set as low as 50, several species of mammals could not be supported in the seven parks.

Re-growing a national park — cow dung with guanacaste seedlings in Costa Rica (*Enterolobium cyclocarpum*). (D.H. Janzen.)

Newmark's findings are not unduly surprising, because none of these parks was set aside expressly to conserve particular mammals. Moreover, parks like Nairobi or Yellowstone (or even Central Park in New York City!) were originally created within a larger area of roughly similar 'wilderness' and so were effectively much larger when they were originally established. Many present-day problems have arisen gradually as settlement and civilization have grown up around such parks. But the magnitude by which these parks today fall short is alarming. Most of the threatened mammals in these western North American parks are top predators, and their disappearance from the parks could trigger further ecological imbalances.

The problem of the scale of natural disturbances is even thornier. How does one go about determining the spatial 'requirements' of a forest fire, a pine beetle infestation, or an epidemic of rinderpest?

Until recently, most natural disturbances were unwelcome inside parks, as they were thought to harm wildlife and scar the landscape. Thus, Yellowstone National Park has a long tradition of suppressing forest fires, even though the lower elevation forests in the Yellowstone region may have burned as frequently as every 20–25 years before the arrival of European settlers[3]. In some areas, aspen stands, an important habitat for some species, are being replaced by conifers, apparently because they are not being rejuvenated by fire. Over the past two decades, park authorites have allowed naturally caused fires in remote areas to burn, provided they do not pose a threat to human life or property.

Disasters

On the other hand, for very rare and localized species or for very small parks, such natural disturbances can be catastrophic. Few conservationists would want to see a fire rage through the world's only stand of Virginia round-leafed birch, however natural such an event might be.

In terms of park design, a further consideration is whether to plan for an 'average' disturbance or a real disaster — the 'fire of the century' or the 'fifty-year flood'. Given the irreversible nature of most land development, it may be wisest to take a rather pessimistic view and to plan for the worst.

There are several ways to make park boundaries more compatible with ecological realities. The simplest solution in the long run is to make the designated parks big enough at the outset. Manu National Park in Peru includes an entire river drainage within its more than 1,600,000 hectares and may be one park that can handle almost any natural disturbance that befalls it. But opportunities to create parks of this magnitude are few and fast disappearing.

A second strategy is to link pre-existing parks, as long recommended by island biogeographers[4-6]. Probably the finest example of this approach is the recent acquisition of La Zona Protectora in Costa Rica by a consortium of conservation organizations, including the Nature Conservancy, the National Park Service of Costa Rica, the National Parks Foundation of Costa Rica, the Organization of Tropical Studies and the World Wildlife Fund. La Zona Protectora is a 7,700-hectare stretch of forest 3–6 kilometres wide that links two important reserves in the country: the upper-elevation Braulio Carillo National Park and the lower-elevation La Selva Biological Station. La Zona Protectora does more than merely increase the total amount of protected land; it attacks the problem of habitat linking the two reserves, it has a much greater additive effect. The combined reserves can support larger populations of many species, and there is now a protected corridor for the approximately 35 species of birds that migrate from the uplands to the lowlands (plus other, less well-known, migratory mammals and insects).

Development

A third strategy is to allow development of the lands surrounding the parks, but to restrict it to those uses and practices that are compatible with conservation goals. Thus, many conservationists in the United States now speak of a greater Yellowstone ecosystem, referring to Yellowstone and Grand Teton national parks (total area 1,022,000 hectares), and parts of the nine national forests (total area 5,747,000 hectares) that encircle them. The management of these national forests is currently the subject of much controversy, as government officials and private citizens debate how activities like logging, mining and grazing should be regulated to avoid harming the flora and fauna of the region.

Another example of an arrangement which is good in principle, but where the actual working is more complicated, occurs in Kenya where the inner 518 km^2 of the Masai Mara Reserve are managed like a park, with no exploitation of the land permitted, whereas in the outer 1,294 km^2 the Masai tribesmen are allowed to graze their cattle[7]. The arrangement sounds fine, but encroachments on the core area are hard to curb. The rules in the Ngorongoro Reserve of Tanzania arguably work better: here the Masai can graze cattle at limited densities anywhere outside the crater, but may only take them into the crater for watering. Despite their problems, restrictive land-use arrangements such as these may be the only means to reconcile the conservation of wide-ranging species with human demands for the land.

Even where development has already harmed much of the land, it is possible to acquire small patches of pristine habitat, along with disturbed areas around it, and then to foster the outward growth of the park into the previously disturbed area. The current scheme to 'grow' a park in dry tropical forest in Costa Rica's Guanacaste Province is an example. The central protected block in Santa Rosa National Park is only about 60 km^2, but the eventual aim is to enlarge the park to several hundred square kilometres by buying low-grade ranch, farm and timber land next to Santa Rosa and on the adjacent Volcan Orosi, and to let the park grow out into it. Only in this way is there likely to be enough area and moist upland refuges to buffer the severely threatened dry forest from normal environmental fluctuations, such as severe droughts and epidemics.

A campaign to buy this land, and create the larger Guanacaste National Park (for a cost of around $10 million) is now underway. If successful, it will demonstrate further that the existing network of parks does not have to be static as long as conservationists continue to come up with imaginative ways to enlarge and improve it. ☐

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544; David S. Wilcove is at The Wilderness Society, 1400 I Street NW, Washington DC 20005, USA.

---

1. Reese, R. *Greater Yellowstone: the National Park and Adjacent Wild Lands* (Montana Magazine, Helena, 1984).
2. Newmark, W. *Biol. Conserv.* **33**, 197 (1985).
3. Norse, E. *et al. Conserving Biological Diversity in Our National Forests* (The Wilderness Society, Washington DC, 1986).
4. Diamond, J. *Biol. Conserv.* **7**, 129 (1975).
5. Wilson, E. & Willis, E. in *Ecology and Evolution of Communities* (eds Cody, M & Diamond, J.) (Harvard University Press, Cambridge, 1975).
6. Harris, L.D. *The Fragmented Forest* (University of Chicago Press, 1984).
7. Williams, J. *A Field Guide to the National Parks of East Africa* (Collins, London 1981).

Superstrings

In search of a smoking gun

Roberto Peccei

SUPERSTRINGS[1] have squarely split the elementary particle community into believers and sceptics. Superstring theories purport to describe a consistent quantum theory of gravity, unified with that of the other known forces, in which the nature of the matter constituents is also specified. No wonder, therefore, that superstring partisans abound. Unfortunately, all this unification occurs in 10-dimensional space-time, so that real physics can emerge from superstrings only as a result of the spontaneous compactification of the extra six spatial dimensions (see the *News and Views* article by Alvaro De Rújula in *Nature* **320**, 678; 1986). These must curl up into a little ball, torus or other compact shape of radius of the order of 10^{-37} cm. Such scales will never be probed experimentally. Thus, say the sceptics, superstrings will forever remain figments in theorists' fevered imaginations. But a new breed of phenomenologists, superstring phenomenologists, is trying to bridge the gap between these two camps, by examining possible patterns left over from the process of compactification.

Although the energy scale of compactification at around 10^{19} GeV is enormous compared with that probed by present-day experiments (around 10^2 GeV), if superstring theories are correct, not very much is supposed to happen in between. Thus, one can hope to extrapolate smoothly through this energy desert, something which was attempted in earlier work on unifying the strong and electroweak forces into a grand unified theory[2]. Also in superstrings, the usual forces are unified in a large symmetry structure before compactification. For the favoured heterotic superstring[3] this is the group $E_8 \times E_8$, where E_8 is the largest exceptional Lie group. In the process of compactification, however, this large symmetry breaks down to a smaller group. Clearly, for these theories to be sensible, this surviving symmetry group should be that connected with the known low-energy forces, or a small extension thereof.

There is another aspect of compactification which is crucial for four-dimensional physics. Superstrings have supersymmetry, a mathematical property which connects matter particles (half-integer spin fermions such as quarks and leptons like the electron) with the 'fields' that cause forces amongst them (integer-spin gauge bosons such as the photon, causing the electromagnetic force). The gauge bosons are associated with the symmetry group structure of the strings, and to each such boson supersymmetry assigns a half-integer spin partner (a fermion). On compactification, only some of these fermion partners survive. The kind and number of those which do are entirely fixed by the geometry of the compact space. If superstrings are to connect to reality, the surviving states must match fermions we see in four dimensions — quarks and leptons — with perhaps a few extra states yet to be observed.

So far there is no proof that the six unwanted spatial dimensions in superstring theories really can curl up, following the string equations of motion, into a space of a given geometry. In fact, it is not clear that if compactification occurs, there is a unique result. (There may be an infinity of possible worlds, of which ours has the charming property that some of its inhabitants may be able to understand it!) However, it is possible to adduce indirect theoretical evidence[4] that, if this compactification takes place, the compact six-dimensional space is a so-called Calabi–Yau space [5,6] or a generalization of these spaces, known as orbifolds[7]. Proceeding on this assumption, Candelas, Horowitz, Strominger and Witten[4] derived two results which brought superstrings much closer to reality. What they found was that the emerging four-dimensional theory had a symmetry group $E_{6'} \times E_8$ where $E_{6'}$ is either E_6 or one of its subgroups, and that the concomitant four-dimensional fermions occurred only in replicas of a representation of E_6 containing 27 particles. This was good news for superstring fans, as E_6 was already known as a nice grand unified theory group in which the low-energy forces fit well. Furthermore, the 15 quarks and leptons of each family fit perfectly within the 27-dimensional representation of E_6. Matter and forces already known match precisely those emerging from this compactification, with a few extra forces and matter states thrown in for good measure.

The trouble is, one can have too much of a good thing. The simplest Calabi–Yau manifolds end up having enormous numbers (of the order of 100) of repetitions of the 27s of quarks and leptons. They also include forces that cause protons to decay very rapidly, totally at variance with observation. A systematic investigation of more favourable manifolds has been carried out by several groups[8–10] in the past year, imposing reasonable boundary conditions to weed out uninteresting manifolds from the start. One of these conditions is that the resulting group, after compactification, should be so small that none of its gauge bosons can lead directly to rapid proton decay. In effect, this requirement necessitates that the E_6 group be broken at compactification. This obtains in manifolds of sufficiently complex topology so that it is possible for some of the E_6 gauge fields to become trapped in holes in the manifold, leading to a symmetry breakdown[11]. Remarkably, precisely for these kinds of manifolds, the Euler characteristic, which details the number of matter replications, is also small. So two potential problems can be alleviated at the same time.

Two generic results emerge from these analyses[8–10]. First, if all the symmetry breakdown occurs at the compactification scale, the surviving symmetry group contains the known forces plus at least one new force, whose associated gauge boson must have a mass comparable to that of the weak gauge bosons of around 100 GeV. Furthermore, the number of families is limited to three, and exotic states in the 27, besides quarks and leptons, survive at low energy. Second, if there are further stages of symmetry breakdown below the compactification scale, it is possible that the low-energy theory contains just the known forces. All exotic matter is, in this picture, superheavy and experimentally inaccessible; and up to four families of quarks and leptons are permitted.

Clearly the first case is more appealing phenomenologically, because it promises that both new forces and matter are there to be found, providing a 'smoking gun' to test the superstring hypothesis. However, this case is not guaranteed and is in itself not without difficulties. For instance, unless certain couplings of the exotic matter states are absent, protons will again decay rapidly. In this sense, the case with intermediate scale breaking is phenomenologically safer, but it is also almost devoid of remnant signals. This is nicely exemplified by the work of Graham Ross and his collaborators[12] at Oxford, who studied in detail one of the few known three-family Calabi–Yau manifolds. At compactification, this manifold either retains the full E_6 symmetry or is broken down to $SU(3) \times SU(3) \times SU(3)$ or $SU(6) \times U(1)$. None of these patterns is nice. However, Ross and his collaborators[12] argue convincingly (although to my mind not totally compellingly) that a sequence of further breakdowns take place which reduce the $SU(3) \times SU(3) \times SU(3)$ theory to precisely $SU(3) \times SU(2) \times U(1)$ — the gauge theory describing the strong and electroweak interactions. Not only are the forces standard at low energy but so is matter, with the exception of two weakly coupled neutral states. What remains as distinct predictions from the original theory are certain relationships between masses and mixing angles of quarks — predictions which are in fair agreement with experiment. Although this analysis is very nice, this is not much of a smoking gun to convince anyone who is not already an *aficionado* of the reality of superstrings.

In my opinion what will be necessary to sway superstring sceptics (among whom I count myself) is some hard evidence for the existence of superpartners to the low-energy excitations we know. Supersymmetry is, after all, a key feature of superstrings and, even after the Calabi–Yau compactification it remains a good symmetry of the four-dimensional theory. Although the matter fermions divorce themselves from the gauge bosons during compactification, they pick up in the process a set of spin zero partners so that, remarkably, a supersymmetry is retained[4]. Eventually, below the compactification scale this supersymmetry must itself break down. How this happens is the least understood piece of the theory, although it is often suggested that the breakdown is triggered by physics associated with the other E_8 group which is unaffected by the compactification. At any rate, this breakdown of supersymmetry is necessary as it is the seed which allows all 'light' masses to be generated in the theory[13], including the mass M_W of the weak gauge bosons. It is therefore natural, and in fact is required by the internal consistency of the theory, that the mass splitting among superpartners generated by this mechanism also be of the order of M_W.

Superpartners with masses of the order of 100 GeV are in the range of discovery for the forthcoming large accelerators at CERN near Geneva, DESY in Hamburg, Fermilab near Chicago and SLAC in Palo Alto. Their discovery, or that of some other debris of the superstring, would provide us with a first glimpse into extra dimensions. ☐

1. Green, M. *Nature* **314**, 409 (1985).
2. Georgi, H., Quinn, H. & Weinberg, S. *Phys. Rev. Lett.* **33**, 452 (1974).
3. Gross, D., Harvey, J., Martinec, E. & Rohm, R. *Phys. Rev. Lett.* **54**, 502 (1985).
4. Candelas, P., Horowitz, G., Strominger, A. & Witten, E. *Nucl. Phys.* **B258** 46 (1985).
5. Calabi, E. in *Algebraic Geometry and Topology* (Princeton Univ. Press, 1957).
6. Yau, S.T. *Proc. natn. Acad. Sci. U.S.A.* **74**, 177 (1978).
7. Dixon, L., Harvey, J.A., Vafa, C. & Witten, E. *Nucl. Phys.* **B261**, 678 (1965).
8. Cecotti, S., Derendinger, J.P., Ferrara, S., Girardello, L. & Roncadelli, M. *Phys. Lett.* **156B**, 318 (1985).
9. Dine, M., Kaplunovsky, V., Nappi, C., Mangano, M. & Seiberg, N. *Nucl. Phys.* **B259**, 549 (1985).
10. Breit, J.D., Ovrut, B. & Segré, G. *Phys. Lett.* **158B**, 33 (1985).
11. Witten, E. *Nucl. Phys.* **B258**, 75 (1985).
12. Greene, B.R., Kirklin, K.H., Miron, P. J. & Ross, G. G. *Phys. Lett.* **B180**, 69 (1986).
13. Ibañez, L. *Phys. Lett.* **118B**, 73 (1982).

Roberto Peccei is at the Deutsches Elektronen Synchrotron (DESY), 85 Notkestrasse, D-2000 Hamburg 52, FRG.

Data bank for carbohydrate structures

THE need for an international data bank for carbohydrate structures was discussed at a recent workshop, sponsored by the US Department of Energy (Aubern, New York, 7–9 August 1986). Such a facility would provide primary structure information analogous to that obtained from the established amino-acid and nucleotide sequence data banks, which would allow the currently impossible systematic search for structural relationships between the 4,000–5,000 carbohydrate structures so far known.

The computer programs required to organize and store the carbohydrate structural data now exist but are more complex than those used for amino-acid sequences. The complexity results from the combinatorial effects of branching in the carbohydrate chain. Conference participants agreed that structural data would be collected by representatives of the different sub-fields of carbohydrate research and organized centrally at the Complex Carbohydrate Research Center, University of Georgia, under the guidance of its director Dr Peter Albersheim.

The estimated costs for the first year of the project are $200,000. A large contribution (40 per cent) will be made by the US Government and the remainder will come from the twelve other participating countries. Users will pay an annual subscription of $100. This improved access to information will be welcomed by both academic and industrial (for example, pharmaceutical, recombinant DNA, oil) concerns.

An international steering committee was set up to coordinate the growth and funding of the project and the data bank should be in full use within a year. The conclusions of the workshop will be available in about two months from the US Department of Energy.

Sarah Hargreaves

Celestial mechanics

Arcs around Neptune

Carl D. Murray

PLANETARY rings now form such a menagerie that it had little need for any new attractions — given the diversity of the ring systems of Jupiter, Saturn and Uranus, with their associated dynamical problems. However, the latest addition to the collection is the planet Neptune. Recent occultation results by Covault *et al.*[1] appear to confirm the presence of incomplete rings or 'arcs' around the planet and a new model by Goldreich *et al.*[2] attempts to provide an explanation for a series of puzzling observations.

Since the fortuitous discovery of the rings of Uranus in 1977 when a star was occulted by ring material, Neptune became an obvious target for studies to detect a ring system. A series of peculiar results began with the claimed detection of a small new satellite by Reitsema *et al.*[3] in 1981. Further occultations failed to detect any ring material (that is, a feature observed before immersion — the disappearance of the star behind the planet — and after emersion — its reappearance) until Hubbard *et al.*[4] observed a single feature from more than one site. As the feature was only observed on one side of the planet they believed that they had detected an 'arc' of material some 13–15 km wide at a radial distance of between 63,000 and 69,000 km.

The occultation on 7 June 1985 (Fig. 1) by Covault *et al.*[1] produced a new twist to the story as the star to be occulted was actually a binary. Although no pre-immersion event was observed there was a single post-emersion event but only one of the binary pair was occulted. Using this information the authors calculated that the feature must be discontinuous over several thousand kilometres. There is no independent confirmation of their result but they estimate a width of 8 km and their accurate orbital radius of the feature also lies within the limits set by Hubbard *et al.* There were no further detections during a subsequent occultation on 25 June 1985. Given the number of negative results to date it now appears that the current observations can be explained by at least one narrow, partial ring composed of one or more arcs covering about 10 per cent of the orbital path. How can such as ring system be maintained?

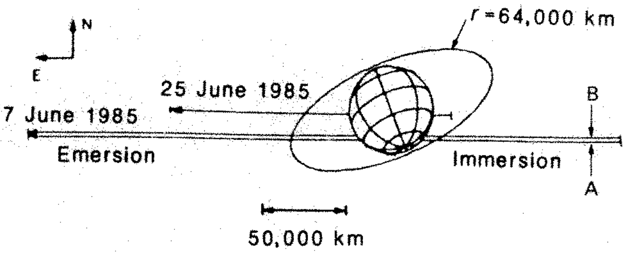


Fig. 1 The apparent paths of stars relative to Neptune for two occultations observed by Covault *et al.*[1] in 1985. Paths A and B refer to the individual stars of the binary pair observed on 7 June. The ellipse shown represents an equatorial radius of 64,000 km, which is a possible location for ring arc occultation events. (Copyright Academic Press.)

Theorists had already considered the dynamical problems posed by arcs in connection with the rings of Uranus. It was originally thought that the outermost ε ring was actually two distinct arcs because occultation features on each side of the planet had different orbital radii. Although it is now known that this is due to the strong ellipticity of the ε ring, the original observations led Dermott and Gold[5] to propose a ring theory involving arcs of material trapped in 'co-rotation' resonances with pairs of Uranian satellites. Resonant phenomena, a common feature in the Solar System, can occur when the periods of two or more objects have a simple ratio and mutual perturbations are enhanced. Lissauer[6] proposed that arcs around Neptune could be explained by the resonant actions of two satellites. The first maintains ring particles in orbits at the classical Lagrange equilibrium points L_4 and L_5 (which precede and trail the satellite by 60°). The second satellite, which lies interior or exterior to the first, provides an additional confinement mechanism to counteract the tendency for the ring to spread due to collisions and other effects.

Goldreich, Tremaine and Borderies have now produced a model for the arc rings of Neptune which relies on the resonant effects of a single satellite in an inclined orbit. The satellite, which may have already been discovered by Reitsema *et al.* in 1981, produces a resonance mechanism which is much more intricate than in previous ring models. They first claim that the arcs lie at one of the strong co-rotation resonances of the satellite. Their location is determined by comparing the orbital periods of a ring particle and the satellite, and also taking into account the longer nodal precession period of the satellite orbit resulting from the oblateness of Neptune. It turns out there will always be an even number of equilibrium positions at the resonant locations and that the ring particles oscillate about each of these positions, forming arcs of material. Goldreich *et al.* then claim that the associated 'Lindblad' resonances involving the orbital periods of the ring particle, the satellite and the pericentre precession period of the ring particle will re-supply the energy lost from the arc through particle collisions. The end result should resemble a series of warped arcs (Fig. 2).

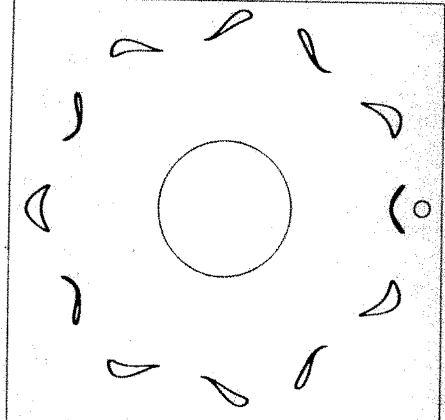


Fig. 2 A possible configuration for ring arcs and a confining satellite around Neptune according to the theory of Goldreich *et al.*[2] Radial variations have been enlarged by a factor of 200. (Copyright American Astronomical Society.)

Given the paucity of hard facts about Neptune's ring system, the model of Goldreich *et al.* seems premature. However, they do make a prediction: the proposed satellite will be found to be in a significantly inclined orbit because otherwise the necessary confining resonances would be too weak. High inclination orbits are the exception rather than the rule for planetary satellites but the most notable exception happens to be Triton, the large satellite of Neptune. The authors speculate that a single event may have produced unusual orbits for all Neptune's satellites.

The study of the dynamics of ring arcs is still in its infancy with many problems still to be tackled. The field will only reach maturity when further observations provide tighter constraints on ring theories. When Voyager 2 reaches Neptune in 1989, the rings of Neptune may be revealed to be even more exotic than we could have imagined. □

1. Covault, C.E., Glass, I.S., French, R.G. & Elliot, J.L. *Icarus* **67,** 126–133 (1986).
2. Goldreich, P., Tremaine, S. & Borderies, N. *Astr. J.* **92,** 490–494 (1986).
3. Reitsema, H.J., Hubbard, W.B., Lebofsky, L.A. & Tholen, D.J. *Science* **215,** 289–291 (1982).
4. Hubbard, W.B., Brahic, A., Sicardy, B., Elicer, L.-R., Roques, F., & Vilas, F. *Nature* **319,** 636–640 (1986).
5. Dermott, S.F. & Gold, T. *Nature* **267,** 590–593 (1977).
6. Lissauer, J.J. *Nature* **318,** 544–545 (1985).

Carl D. Murray is in the Theoretical Astronomy Unit, Queen Mary College, University of London, London E1 4NS, UK.

Optical physics

Growing expectations from squeezed states of light

Dan Walls

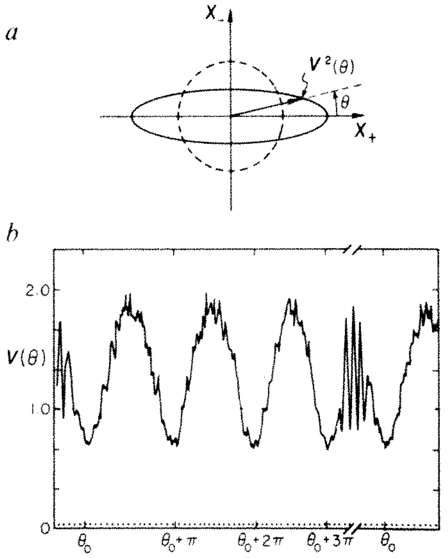
THE past year has seen remarkable progress in the experimental generation of squeezed states of light, which previously were known only to theorists. Squeezed states have the property that the quantum fluctuations in one phase of the wave motion are less than those of the vacuum state of the electromagnetic field, although the fluctuations in the quadrature phase (90° out of phase with the other) compensate by being greater than in the vacuum. In this sense, in one phase squeezed states may beat the quantum limit imposed by the individual nature of the photons, the 'shot noise' of a light beam. There have recently been experimental reports of squeezed states of light generated by four-wave mixing in atomic vapour and optical fibres, and optical parametric oscillation. In the latter, a reduction in quantum fluctuations of more than a factor of two below the vacuum level has been achieved.

The interest in squeezed states of light arises from their possible application in gravitational wave detectors and optical gyroscopes, where they would enhance the sensitivity of interferometric measurements of phase. The reduced fluctuations in one phase offer the possibility of improving the signal-to-noise ratio in quantum-limited optical communications systems. New kinds of experiments in atomic spectroscopy are also feasible. (See my review of the general properties of squeezed states in *Nature* **306**, 141; 1983.)

Squeezed states are generated by nonlinear optical processes involving phase dependent amplification and deamplification of the phases in quadrature. Examples of such processes include four-wave mixing and optical parametric oscillation. The recent experiments examine different variations of these systems.

The first experimental results were reported one year ago by the group of the AT&T Bell Laboratories (Slusher, R.E., Hollberg, L.W., Yurke, B., Mertz, J.C. & Valley, J.F. *Phys. Rev. Lett.* **55**, 2409; 1985) in an experiment in a beam of atomic sodium. The beam is pumped with a coherent dye laser at 589 nm. To achieve a sufficient nonlinear interaction to squeeze the vacuum fluctuations, the laser frequency should be tuned near an atomic resonance. However, as spontaneous emission will destroy the squeezing, the pump laser is detuned from the atomic resonance by 300 atomic linewidths. The vacuum field in an optical cavity noncollinear with the pump laser is squeezed by the four-wave mixing process, and the fields in cavity modes symmetrically displaced from the pump frequency (sidemodes) by 80 atomic linewidths can be detected. This non-degenerate four-wave mixing process ensures the sidemodes are outside the spontaneous emission profile of the atom.

The quantum fluctuations in the output at the two cavity sidemodes are detected by beating with part of the laser which acts as a local oscillator. The local oscillator phase is varied with respect to the sidemodes to detect the quadrature with least fluctuations. The noise reduction observed in this experiment is 17 per cent below the vacuum level, in agreement with theoretical estimates for the parameters of the experiment. But theoretical calculations predict that squeezing should yield a factor of ten reduction in noise at higher pump powers. Another experiment using atomic sodium in a forward geometry in the Doppler broadened atomic vapour regime and no cavity has been performed at the Massachusetts Institute of Technology (Maeda, M.W., Kumar, P. & Shapiro, J.H., personal communications). This experiment has to date achieved a squeezing level of 4 per cent below the vacuum. A further experiment with a sodium atomic beam investigates a new regime for which atomic and cavity timescales are comparable. Initial experiments at the University of Texas in Austin have obtained noise reductions of 13 per cent below the vacuum (Orozco, L.A., Raizen, M.G., Xiao, M. & Kimble, H.J., personal communication).



a, Phase plot of the uncertainties in the quadrature amplitudes of the electric field. The solid line represents the variance $V^2(\theta)$ of the field $X(\theta) = X_+ \cos\theta + X_- \sin\theta$ as a function of θ for a squeezed state; the dashed line is for vacuum fluctuations. *b*, Measurement of the phase dependence of the quantum fluctuations (noise power) in a squeezed state produced by degenerate parametric down conversion. The plot corresponds roughly to the quantity $V(\theta)$ as in *a*. The minima represent a 50 per cent reduction in noise power relative to the vacuum noise level. Note the ordinate is a linear scale in noise voltage. (Courtesy of H.J. Kimble.)

A four-wave mixing experiment but in a completely different type of material, an optical fibre, has been performed by a group at the IBM Laboratory in San José (Shelby, R.M., Levenson, M.D., Perlmutter, S.H., De Voe, R.G. & Walls, D.F. *Phys. Rev. Lett.* **57**, 691; 1986). Light from a krypton ion laser at 647 nm is transmitted through 114 m of glass fibre. Because the laser frequency is far away from any absorption level in the glass fibre, spontaneous emission is not a problem, but Brillouin scattering from acoustic phonons in the fibre introduces excess noise. The fibre is cooled to 4 K with liquid helium to reduce this scattering. The long interaction length in the fibre together with the high intensity of light focused in the fibre compensates for the rather weak optical nonlinearity of glass. Squeezing occurs in continuous pairs of sidemodes symmetrically displaced from the laser frequency. A heterodyne detection scheme is achieved by beating with the original strong laser beam which acts as a local oscillator and can be phase-shifted with respect to the sidebands. This system offers in principle squeezing over a broad bandwidth. However, mainly because of the excess noise from the Brillouin scattering the squeezing observed is limited to 12 per cent.

An experiment using a very different approach to the problem of squeezed state generation has been followed by a group at the University of Texas in Austin (Wu, L.A., Kimble, H.J., Hall, J.L., & Wu, H. *Phys. Rev. Lett.*, in the press) with spectacular success (see figure). The process of parametric down conversion, in which photons of frequency 2*ω* split to produce correlated pairs of photons at the subharmonic frequency *ω*, is used to generate 63 per cent squeezing.

The experiments are perfomed using a degenerate parametric oscillator consisting of a lithium niobate crystal inside a resonant cavity. The cavity is driven by light at the harmonic frequency (0.53 nm) by frequency doubling a Nd-Yag laser (1.06 nm). The down-converted light emerging from the cavity is combined with the original laser emission (at 1.06 nm) using a balanced homodyne detector to observe the squeezing. Although squeezing by more than a factor of two is observed, a quantitative comparison with theory suggests that in fact the intracavity field of the optical parametric oscillator would produce more than ten times squeezing in the absence of losses in the current experiment. Kimble estimates that with improvements in detector efficiencies and cavity parameters the squeezing that can be observed will approach this milestone.

At this level many of the proposed applications of squeezed light look promising. Progress in this area has so far exceeded expectations and we can reasonably expect further exciting developments in the next year. ☐

Dan Walls is Professor of Physics at the University of Waikato, Private Bag, Hamilton, New Zealand.

Plant biochemistry

Discovery of the most abundant inhibitor in the world?

Christine Foyer

The regulation of the activity of the primary enzyme of photosynthetic carbon assimilation, ribulose-1,5-bisphosphate carboxylase, has remained one of the most confusing and elusive areas of plant biochemistry but the recent discovery and identification of an endogenous nocturnal inhibitor, namely 2-carboxy-D-arabinitol-1-phosphate (Gutteridge, S. *et al. Nature* **324,** 274; 1986 and Berry, J. A. *et al. Proc. natn. Acad. Sci. U.S.A.*, in the press), should help resolve discrepancies in our understanding of the modulation of what must be one of the most important enzymes in the biosphere. Together with related observations of a light-dependant activation protein that facilitates optimum enzyme function at ambient levels of carbon dioxide (Salvucci, M.E., Portis, A.R. and Ogren, W.L. *Photosynth. Res.* **7,** 193-201; 1985), this work forms the basis for a new understanding of the overall regulation of carbon assimilation.

Ribulose-1,5-bisphosphate carboxylase was first identified in 1956 but it was not until the mid 1970s that sufficient activity could be achieved *in vitro* to account for the rates of photosynthesis found *in vivo*. Activation of the enzyme requires carbamate formation on a lysine residue of the large subunit followed by reaction with Mg^{2+} ions to form a catalytically active enzyme–CO_2–Mg^{2+} complex but this knowledge provides only a glimpse of the functional capacity and regulation of the enzyme *in vivo*. Full activation of the enzyme *in vitro* requires high levels of carbon dioxide yet it occurs in leaves at ambient CO_2 concentrations. A further problem is that low concentrations of the enzyme activated with high CO_2 and high Mg^{2+} catalyse carboxylation showing distinctly biphasic kinetics with high rates of activity that decline after the first 15–30 seconds of the reaction to a lower steady rate.

A diurnal fluctuation in the specific activity of ribulose-1,5-bisphosphate carboxylase extracted from leaves has been observed in many plant species. Similarly, light activation and dark deactivation of the enzyme have been demonstrated in leaves, leaf protoplasts and isolated chloroplasts. The activation state of ribulose-1,5-bisphosphate carboxylase in leaves can be correlated with irradiance in a manner that cannot be attributed to changes in stromal *p*H and Mg^{2+}. A dark-induced inhibitor of the enzyme may be responsible for this modulation. When leaves are illuminated, photosynthetic carbon fixation shows an induction phase (or lag period) of several minutes before maximum rates are attained. Previous results have suggested that light activation of ribulose-1,5-bisphosphate carboxylase is not an important factor in determining the duration of the induction period but the presence of 2-carboxy-D-arabinitol-1-phosphate adds an extra factor, at least in species where the inhibitor is abundant.

The presence of a dark-induced inhibitor of ribulose-1,5-bisphosphate carboxylase is at once intriguing and puzzling. The degree of activation of the enzyme found in darkness is highly variable and it is not clear why photosynthetic carbon assimilation requires that ribulose-1,5-biophosphate carboxylase is relatively inactive in darkness. Several light-modulated enzymes are present in the pathway whose activity is negligible in the dark. These are activated on illumination long before photosynthesis overcomes the induction period. Now we find that there is another site of inhibition, and nature obviously requires a complete shutdown of the reductive pentose phosphate pathway in darkness.

2-carboxy-D-arabinitol-1-phosphate has so far only been identified in a limited number of species. There may, however be interspecific variation in the metabolism of the inhibitor and also differences in the conditions required for its retention during extraction procedures. We now need to know the biochemical pathways of synthesis and metabolism of this inhibitor and their regulation. Dr A. J. Keys (Rothamsted Experimental Station) has suggested that hamamelose bisphosphate could be a precursor of the inhibitor especially as this sugar-bisphosphate is made and rapidly turned over in the chloroplast. Synthesis and degradation may well be related to irradiance and therefore participate in the regulation of photosynthesis.

Ribulose-1,5-bisphosphate carboxylase has justly been called the most abundant protein in the world. As equimolar amounts of 2-carboxy-D-arabinitol-1-phosphate are required for optimum inhibition, this metabolite must be present in abundance in darkened leaves to be effective. This feature certainly limits any value this compound may have as a potential herbicide. But an understanding of the biochemistry and functional role of this inhibitor may suggest new areas for genetic manipulation of photosynthesis. ☐

Christine Foyer is at the Research Institute for Photosynthesis, University of Sheffield, Sheffield S10 2TN, UK.

---

100 years ago

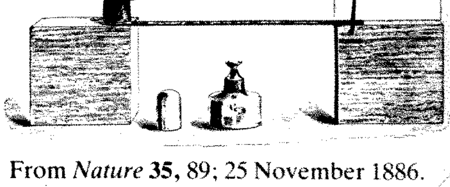
I venture to call attention to a simple and effective way of demonstrating the linear expansion of solids when heated, first suggested, I believe, by M. Kapoustine (*Journal de Physique*, December 1883, p. 576). The principle is, to magnify the slight extension of a bar by causing the end of it to roll upon a needle, and thus turn the latter round and move a pointer attached to it through a sensible arc.

A small flat rod of the material to be examine is laid upon two wooden blocks, placed about 25 cm apart. A weight is put upon one end of the rod to keep it from moving; under the other end, at right angles to the length of the rod, is laid a fine sewing-needle, to the eye-end of which a light pointer of straw, about 16 or 20 cm long, is attached by sealing-wax. Behind the pointer (which is painted black) a screen of white cardboard is fixed on the wooden block by drawing pins.

Before the experiment is shown to an audience, it is well to make sure that the needle rolls fairly and freely between the bar and the block. Such precautions, however, are not in the slightest degree necessary for school-work; for there is always one thing which gives the typical boy greater pleasure than to see an experiment succeed, and that is — to see it fail.

H. G. MADAN

Eton College



From *Nature* **35,** 89; 25 November 1886.

Community ecology

Patterns in three dimensions

Paul H. Harvey and John H. Lawton

MUCH of the best in ecology comes from recognizing and correctly interpreting patterns of community structure. Occasionally, the interpretation of one pattern can throw light on another, even in an apparently disparate area of the science. Just such a claim is made by Brown and Maurer on page 248 of this issue[1]. They find that within a group of species that exploit the same class of environmental resources in a similar way, larger-bodied species seem to use up a disproportionate amount of the resources available. They go on to argue that, if their results are correct, a mixture of individual and species selection could provide an explanation for Cope's rule—the finding that members of taxonomic groups typically become larger sized over evolutionary time[2].

Brown and Maurer's analyses concern the relationship between population density and body size in assemblages of birds, fishes, rodents and plants. Larger bodied species are found at lower population densites. Damuth earlier pointed out that population density among African ungulates decreases with approximately the -0.75 power of body size[3]. Because mammalian basal metabolic needs increase with the 0.75 power of body size[4], it seemed to follow that populations of large and small species were using about the same amount of energy. Peters[5] subsequently claimed that the exponents relating population density to increases in body size over a wider range of animal groups are less than -0.75, probably nearer -1.0, indicating that smaller bodied species use more resources. In contrast, the latest analyses by Brown and Maurer suggest that across species exploiting similar resources in a similar way within communities, exponents are appreciably shallower than -0.75. Hence, large species seem to use more energy per population than do small species.

Population density and body size can be considered with a third community variable, the number of species. Each of the three pairs of variables — population density and body size[3-5]; population density and number of species[6-8]; and body size and number of species[9] — are generally considered independently of each other. But knowledge of any two of these relationships automatically provides information about the third (J.H.L. & S.C. Pimm, in preparation). Indeed, if we understand the ecological and evolutionary determinants of any two of the patterns, then the third follows as a necessary consequence. With one exception[9], theories that have been developed to explain each separate relationship in this three-dimensional space have not recognized their interdependence. Even if the energy flow per species is lower in small organisms, because there are almost always more small species than large species in communities[10], the proportion of total community energy flow through small species may nevertheless exceed community energy flow through large species.

The problem then is to explain why energy, or resources in general, are partitioned by many small, or a few large species, and how shares are allocated[10]. Brown and Maurer's observations on population energy use are an important step in this direction. Ecologists need to pay much more attention to establishing reliable empirical relationships between body size, population abundance and number of species in plant and animal communities, as well as to their theoretical interrelationships.

An example comes from the study of species abundance curves that describe the relationship between population density and the number of species. For many years it has been recognized that frequency histograms of the relative abundances (population densities) of the different species exploiting similar resources, when plotted with the abundance on a logarithmic scale, are normally distributed. The distribution is said to be log normal. Futhermore, the number of species in the sample often seems to bear a particular relationship to the variance of the distribution[10]. Log normal distributions with this property are termed 'canonical'. Models of the ways resources might be partitioned among the different species have been produced, the most notable being MacArthur's broken stick[11] and Sugihara's sequential niche division[8] models. The latter actually predicts a species abundance distribution very similar to that observed (it is both log normal and canonical). But such models rest on the assumption that individuals of the different species have identical resource requirements. As Brown and Maurer and others have shown, this assumption does not hold and, furthermore, there is usually an inverse relationship between individual resource requirements and population density. One consequence of the distribution of relative species abundances is that energy may be partitioned among species according to a log normal distribution, but that distribution will not be canonical[12]. It may be that, for once, ecologists have less to explain than they originally thought. ☐

1. Brown, J.H. & Maurer, B.A. *Nature* **324**, 248 (1986).
2. Stanley, S.M. *Evolution* **27**, 1. (1973).
3. Damuth, J. *Nature* **290**, 699. (1981).
4. Kleiber, M. *The Fire of Life* (Wiley, New York, 1961).
5. Peters, R.H. *The Ecological Implications of Body Size* (Cambridge University Press, 1983).
6. Preston, F.W. *Ecology* **29**, 254 (1948).
7. May, R.M. in *Ecology and Evolution of Communities* (eds Cody, M.L. & Diamond, J.M.) 81 (Harvard University Press, 1975).
8. Sugihara, G. *Am.Nat.* **116**, 770 (1980).
9. May, R.M. in *Diversity of Insect Faunas* (eds Mound, L.A. & Waloff, N.) 188 (Blackwell, Oxford, 1978).
10. Preston, F.W. *Ecology* **43**, 185; 410 (1962).
11. MacArthur, R.H. *Am.Nat.* **94**, 25 (1960).
12. Harvey, P.H. & Godfray, H.C.J. *Am. Nat.* (in the press).

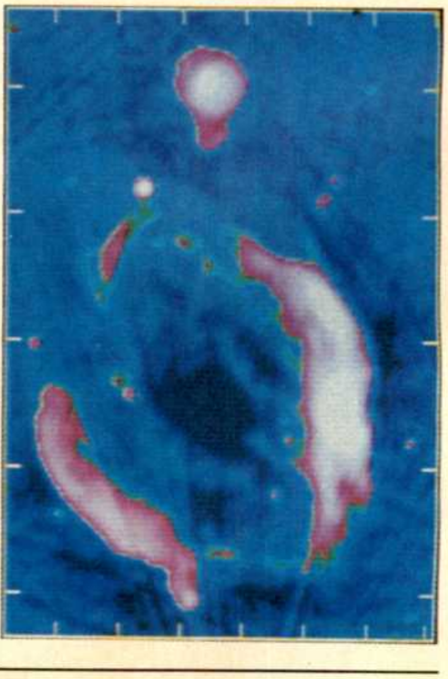
Paul H. Harvey is in the Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, and John H. Lawton is in the Department of Biology at the University of York, Heslington, York YO1 5DD, UK.

A new nebula

A COLOUR-ENHANCED image of the Circinus X-1 region of the galactic plane obtained with the Sydney University Molonglo Observatory telescope at 843 MHz, with Circinus X-1 itself in the upper region of the image.

Circinus-X-1 may well be one of the more peculiar cosmic objects. It appears to be a binary system, one part of which is so compact it may be a neutron star or a black hole.

As discussed by R.F. Haynes *et al.* on page 233 of this issue, the association of Circinus X-1 with with newly discovered radio nebula is in contrast to most X-ray emitting binary stars. Circinus X-1 is in some ways similar to Cygnus X-1, but no analogous extended radio component has been discovered associated with Cygnus X-1. Comparison is also made with SS433 and the Crab nebula. It is suggested that the nebula results from accumulated energetic particles emitted during flaring activity.



Solar System

How fast does Halley spin?

Philip Campbell

Astonishingly, the rotation rate of the nucleus of comet Halley has suddenly become the focus of intense debate, eight months after the famous spacecraft encounters with the comet. The fact that significant strides have been achieved in the analysis of spacecraft and ground-based photometry and spectroscopy was clear from a recent meeting*. But above all, many participants still seemed to be recovering from the announcement at the previous week's Halley symposium in Heidelberg of the detection of a period in light curves more than three times longer than that hitherto adopted for the nuclear rotation.

The simplest way to investigate cometary rotation, given the acute difficulty in resolving cometary nuclei from Earth, is to seek regular variations in brightness. Because of the vagaries of the allocation of telescope time, there are no light curves available that are continuous since the comet's first detection, making an unbiased Fourier or maximum entropy analysis impractical. One extended collection of data was discussed by M. J. S. Belton (Kitt Peak) but, in reporting a period of 54 hours, he emphasized that he was guided by the selection of possible periods by two previous results — observations from the Japanese spacecraft Suisei of a 2.2-day (53-hour) variability in the ultraviolet intensity of the vast hydrogen cloud surrounding Halley, and an analysis by Z. Sekanina and S. Larson of 1910 dust jet images.

The more sophisticated way to measure the rotation rate of a nucleus is to send three spacecraft to look at it. On 6 March, Vega 1 saw the long axis of the nucleus pointing in a Sunward direction. Three days later, Vega 2 passed Sunward of the nucleus and obtained a broadside image. The two images were consistent with a 480 degree rotation about a spin axis perpendicular both to the long axis of the nucleus and to the orbital plane, as would accord with a 53-hour period. More recently the Giotto and Vega images were considered in combination by both imaging teams who independently arrived at a 53-hour period.

For much of the Heidelberg meeting the cometary scientists had been happy with this rotatory paradigm. But the harmony was broken somewhat when R. L. Millis and co-workers from the Lowell Observatory announced the detection of a strong 7.4-day period in ground-based photometric observations using emission bands such as C_2, CN and OH (*Nature*, in the press). The light curves, measured in March and April this year, reveal a characteristic double-peaked structure which Millis identified as that produced by active regions entering sunlight once every rotation period.

No-one questions the reality of the 7.4-day period in the data — it has now been reported by other observers. For example M. F. A'Hearn *et al.* (Maryland), following their unexpected discovery of the existence of spiral jets of CN thought to be the photodissociation product of dust emissions (*Nature*, in the press), have found that a 7.4-day period and not a 2.2-day period satisfies the relative orientations of the jets at different times. In contrast, B. Smith (Arizona) of the Vega imaging team is adamant that the orientations of the nucleus at the three spacecraft encounters cannot accommodate such a period. It is all the more intriguing, therefore, that another member of the Vega team, J. L. Bertaux (Service d'Aeronomie, Verrières-le-Buisson) is less categorical. Nevertheless it does seem that proponents of the 7.4-day rotation period, if they are to conform to the imaging constraints, will have to appeal to a rotation axis more inclined to the long axis of the nucleus and a more complicated nuclear geometry than has hitherto been thought necessary.

As it happens there is still scope for further image enhancement. If the 2.2-day rotation period survives such scrutiny, what can be made of the 7.4-day light curve variations? Their regularity points to some dynamical origin and that most widely assumed seems to be a free wobble (a motion not maintained by a torque) equivalent to the terrestrial Chandler wobble of the Earth, though much closer to the rotation rate of the comet than is the rate of terrestial Chandler wobble to the rotation rate of the Earth. At Heidelberg, K. Wilhelm of the Max Planck Institute, Lindau, who had analysed the motion of an appropriate-sized ellipsoid, reported that a wobble period of about six days was not unreasonable and was consistent with the spacecraft images.

Despite the fact that such a wobble could not be observed directly from outside the cometary coma, its behaviour, if manifested in photometric light curves, could yield useful information about the interior of the nucleus. The wobble would presumably be excited by the jet activity as the comet passed through the vicinity of perihelion. Over the subsequent orbital period, as shown by Wilhelm, it should be damped out by internal friction. But in Paris, M. Festou (Observatoire de Besançon) presented an amalgamation of all available photometric data taken between January 1984 and February 1985, showing that a 7.4-day period was significant throughout. As J. Lissauer (U. C. Santa Barbara) emphasized, interpreting this as a wobble would imply a significant amplitude near the very onset of emissions from the nuclei, in turn implying small internal damping and a comparatively rigid nuclear interior.

Although such detailed studies of the behaviour of Halley have proved unexpectedly engrossing — and may remain so for some time — more wide-ranging questions were also addressed in Paris, relating in particular to the vast but putative Oort cloud of comets from which Halley is thought to have originated. The darkness of cometary nuclei, emphasized by spacecraft images of Halley, combined with the new findings of the shape and degree of activity of the nucleus, allowed P. Weissman (Jet Propulsion Laboratory) to generalize and estimate afresh the total mass of the cloud, increasing it over a thousandfold above Oort's original estimate of a tenth of an Earth mass or less. This could imply that more comets have collided with Earth than hitherto thought, also strengthening the idea that an early generation of comet-like objects may have delivered the stuff from which the atmospheres and oceans to be found on the terrestrial planets are derived.

The perturbations of the Oort cloud thought to produce long-period comets have traditionally been ascribed to random encounters with passing stars. Recent theoretical work, however, indicates that ever-present galactic tidal forces should also play a significant role. A. H. Delseume and M. Patmiou (Toledo) plotting the aphelia of 152 selected long-period comets in galactic coordinates, reported 'zones of avoidance' at the Galactic poles and equator, consistent with a tidal effect.

In reviewing the progress in the analysis of spacecraft data, E. Grun (Max Planck Institute, Heidelberg) described surprising results from the PUMA dust impact mass analysers on the Vega spacecraft. Small particles ($< 1\ \mu m$) have been found with relative abundances of C, N and O very similar to those in the Sun and much higher than those found in carbonaceous chondritic meteorites. Similarly, gaseous isotope ratios measured from spacecraft and from the ground are consistent with solar abundances. Though the spacecraft results are still preliminary, all the signs are that — as has always been hoped — comets contain by far the most primitive and pristine material to be found in the Solar System. □

Philip Campbell is Physical Sciences Editor of Nature

---

*Annual meeting of the Division of Planetary Sciences of the American Astronomical Society, Paris, 4–7 November 1986.

Yeast genetics

Genetic control mechanisms: transcriptional twisting

Ian W Dawes

YEAST geneticists and molecular biologists have long been advocating yeast as a model eukaryotic organism as well as an organism for exploitation by the biotechnologist Reservations about the usefulness of such a model are based largely on the relatively uncomplicated nature of the yeast life cycle and partly on differences from higher eukaryotes in the mitotic and meiotic processes At a recent meeting*, however, it became clear that much progress has been made in the study of those cellular processes for which yeast is ideally suited One example of such a process is the regulation of transcription

Now that important regulatory sequences have been identified upstream of several genes in yeast, the search is on in earnest to find proteins binding to these sequences and to determine how this binding leads to specific control as well as initiation of transcription With regard to the nature of the upstream sequences concerned with positive control, the seven genes for allantoin degradation (*DAL1* to *DAL5, DAL7, DUR1,2*) which are all under the control of the *DAL81* gene, contain two types of consensus region (T G Cooper, University of Tennessee, Memphis) For efficient constitutive expression, Cooper proposed an upstream expression sequence (UES) which is found in front of all the above genes at least once For the highly regulated *DAL4, DAL7* and *DUR1,2* genes there is a second sequence, an upstream modulation sequence (UMS) present in multiple copies, which is involved directly in the inducibility of these genes For the less inducible genes (*DAL1* and *DAL2*) there are fewer UMS elements, and these exhibit only limited homology with the consensus region For the constitutive genes (*DAL3* and *DAL5*) there is no UMS

The distinction between expression and regulation sequences is not radically different from what has been reported for other genes that have been examined (and maybe soon the terminology will be agreed), but the nice point of this work is the extent to which the correlation between control sequences and inducibility holds for a group of genes under one type of control

It also illustrates the extent to which UMS and UES sites can be fairly scattered around in the 5′ upstream region of a naturally regulated system For negative control of the *CAR1* gene, Cooper's group has found a 13-base pair sequence (an upstream repression sequence, URS) that is necessary for repression, and which can control *CYC1* expression if cloned upstream or downstream of the upstream activating sequence (UAS) elements This sequence is present in the 5′ flanking regions of more than a dozen genes of widely diverse functions, including nitrogen catabolism, protein synthesis and glycolysis, and it may be part of a more global regulatory system, or possibly a generic sequence associated with a repressor binding site as it does show some homology with the phage λ repressor binding site

Several points have emerged about proteins binding at the UAS elements The DNA-binding proteins that have been sequenced have short domains for DNA binding relative to the overall length of the molecules (for example, the amino-terminal 74 residues from 881 for the product of the regulatory *GAL4* gene (G Gill, Biochemistry and Molecular Biology, Harvard University) and 60 carboxy-terminal residues from 282 for the *GCN4* gene product conferring general amino-acid regulation (K Struhl, Harvard Medical School) These DNA-binding regions show the helix-turn-helix motif of prokaryotic repressors Other regions of these control proteins are necessary for the activation of transcription, thus fusions of the carboxy-terminal end of the *GAL4* gene product and the *Escherichia coli lexA* protein will bind to *lexA* promoters and initiate transcription in a heterologous yeast system (Gill) The elegant work from Struhl's laboratory with the *GCN4* gene product shows that in addition to the short DNA binding site at the carboxy terminus, this protein has a central and even shorter (19-residue) acidic domain needed for the activation function There are similar acidic regions in other yeast regulatory proteins and in the HMG proteins thought to be important in regulation of higher cell development (Struhl) Deletions between the DNA-binding and acidic domains do not abolish the activity of the protein *in vivo*

The presence of a single DNA-binding regulatory protein at the UAS is not sufficient for control — other proteins are involved This has been found for the upstream control of the two cytochrome *c* genes *CYC1* and *CYC7* (L Guarente, Massachusetts Institute of Technology) *CYC1* expression is regulated by haem and glucose from two elements, *UAS1* and *UAS2* Dissection of the 73-base pair *UAS1* region reveals one protein (plus) binding the one region, and two others (HAP1 and RC2) to another How do the proteins binding at these UAS sites affect the initiation of transcription which can be located from 250 to 1,000 or more base pairs downstream from the UAS site? Rudi Planta (Vrije Universiteit, Amsterdam) described the proponents of various hypotheses as being either "twisters, sliders, oozers or loopers" (see Ptashne, M *Nature* **322,** 697–701, 1986) Gill, a confirmed looper, suggested for the *GAL1* gene that up to four molecules of the *GAL4* product bind at UAS_G and interact with other proteins binding nearer the start of transcription, thereby introducing a loop in the DNA and bringing the UAS and the transcription start site into juxtaposition

With all the current discussion of whether the human genome should be completely mapped and sequenced, and who will pay for the exercise, it is interesting to see how far similar exercises have progressed in an organism with only a fraction of the haploid DNA content M Olson (Washington University School of Medicine, St Louis) reported on progress in the task of restriction-mapping the entire yeast genome at the 2-kilobase level By using 'bottom up' screening of about 5,000 λ inserts for overlaps, about 600 local maps have been identified, which should reduce to about 200 on a second pass through the restriction data The other approach of 'top down' mapping using OFAGE electrophoresis to separate the 16 yeast chromosomes and map them individually with the rare cutting restriction enzymes *Not*I and *Sfi*I, generating about 200-kilobase fragments, is also well under way, with more than half of the 55 *Sfi*I fragments mapped, and complete maps for eight of the chromosomes One of the early goals of this work should soon be realized the production of an ordered set of λ clones to allow precise location of a newly cloned gene in a single hybridization experiment The organization of chromosome 1 (the smallest chromosome, in which 250 kilobases can encode only about 100 genes) is the furthest advanced Kaback and colleagues (New Jersey Medical School, Newark) have mapped 160 kilobases of almost contiguous DNA representing 70 per cent of the chromosome, and have also studied transcription from genes previously undetected by mutation

Any report of a meeting of the size and interest of this one is bound to be fragmentary and selective Fortunately, for the first time the complete abstracts of a meeting in this series are available in published form as a supplement to the journal *Yeast* □

Ian W Dawes is in the Department of Microbiology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, UK

---

*13th conference on yeast genetics and molecular biology held at Banff Alberta 1–5 September 1986

Fingers and helices

SIR—On several occasions the interaction between *Xenopus* transcription factor IIIA (TFIIIA) and 5S RNA genes has been discussed in *Nature*, News and Views[1-3] TFIIIA is a good example of the so-called finger proteins[4], eukaryotic regulatory proteins which contain variable numbers of a characteristic tendemly repeating motif of about 30 amino acids[5,6] It is currently thought that these sequence repeats represent small nucleic acid binding domains, each folded about a tetrahedrally coordinated zinc ion The coordination is provided by appropriate side chain atoms from fully conserved histidines and cysteines This had led to a model in which TFIIIA is essentially a 'string of beads' that zigzags along one side of a nucleic acid double helix in an extended fashion[7]

Our observation[5] of the repeating motif in TFIIIA, which has unfortunately been overlooked by most authors, allowed us to make a reasonable prediction of the average secondary structure of the repeat We have now extended the secondary structure analysis to include all the finger sequences known to us (Table 1) The prediction technique[8] involves calculation of smoothed plots for each of the three conformational preference parameters (α-helix, β-strand and turn) for individual sequence repeats The smoothed curves are then averaged over all the repeats The most strongly predicted region spans alignment positions 17–23 (Table 1) where the mean smoothed helix preference varies from 1 1 to 1 15 Given that 1 0 indicates a neutral preference and that 1 4 is the largest preference, the helical prediction over the 39 repeats is clear

Table 1 Aligned 'finger' sequences

*Repeat	1 10 20 30	*Repeat	1 10 20 30
1	YICSFADCGAAYNKNWKLQAHLC-KHTGEKP-	24	HICPI--CGVIRRDEEYLELHMN-LHEGKTE-
2	FPCKEEGCEKGFTSLHHLTRHSL-THTGEKN-	25	KQCRY--CPKSFSRPVNTLRHMR-SHWDKKK-
3	FTCDSDGCDLRFTTKANMKKHFNRFHNIKICV	26	YQC--EKCGLRFSQDNLLYNHRL-RHEAEENP
4	YVCHFENCGKAFKKHNQLKVHQF-SHTQQLP-	27	IICSI--CNVSFKSRKTFNHHTL-IHKENRPR
5	YECPHEGCDKRFSLPSRLKRHEK-VH---AGY	28	HYCSV--CPKSFTERYTLKMHMK-THEGDVVY
6	PCKKDDSCSFVGKTWTLYLKHVAECH-QDL--	29	F-CLI--CNTTFENKKELEHHLQFDH--DVS-
7	AVC--DVCNRKFRHKDYLRDHQK-THEKERTV	30	LHCRR--CRTQFSRRSKLHIHQKLRCGQDFSV
8	YLCPRDGCDRSYTTAFNLRSHIQSFHEEQRP-	31	FVC--EVCTRAFARQEHLKRHYR-SHTNEKP-
9	FVCEHAGCGKCFAMKKSLERHSV-VHDPEKRK	32	YPCGL--CNRCFTRRDLLIRHAQKIHSGNLGE
10	FTCKI--CSRSFGYKHVLQNHER-THTGEKP-	33	FVC--NYCDKTFSFKSLLVSHKR-IHTGEKP-
11	FECPE--CDKRFTRDHHLKTHMR-LHTGEKP-	34	YEC--DVCQKTFSHKANLIKHQR-IHTGEKP-
12	YHCSH--CDRQFVQVANLRRHLR-VHTGERP-	35	FECPE--CGKAFTHQSNLIVHQR-AHMEKKP-
13	YTC--EICDGKFSDSNQLKSHML-VHTGEKP-	36	YGCSE--CGKTFAQKFELTTHQR-IHTGERP-
14	FEC--ERCHMKFRRRHHLMNHKCGIQSPPTPA	37	YEC--NECAKTFFKKSNLIIHQK-IHTGEKR-
15	FEC--EFCHKLFSVKEDLQVHRR-IHTKERP-	38	YECSE--CGKSFIQNSQLIIHRR-THTGEKP-
16	YKC--DVCGRAFEHSGKLHRHMR-IHTGERP-	39	YECTE--CGKTFSQRSTLRLHLR-IHTGEKP-
17	IPCHI--CGEMFSSQEVLERHIKADTCQKSEQ		ααααααα
18	ATC--NVCGLKVKDDEVLDLHMN-LHEGKTE-		
19	LECRY--CDKKFSHKRNVLRHME-VHWDKKK-		
20	YQC--DKCGERFSLSWLMYNHLM-RHDAEENA	†	h q k fs k v k a spqkp
21	LIC--EVCHQQFKTKRTYKHHLR-THQTDRPR		f-C--n-C-r-yt-r--l--H-r-iHtgnrg-
22	YPCPD--CEKSFVDKYTLKVHKR-VHQPVEKP		y e v e
23	QECTT--CGKVYNSWYQLQKHISEEHSKQ-PN		d l d

*Repeats 1 – 9 are from *Xenopus* TFIIIA 10 – 14 are from *Drosophila Kt*, 15 – 16 from *Kr*h (R Schuh and H Kackle), 17 – 22 from beta *sry*, 23 – 29 from detla *sry* and 30 from IB142 (R Baldarelli and J Lengvel) 31 – 32 are from yeast ADRI and 33 – 39 from mouse MKr3 (K Chowdhury and P Gruss)

†The consensus for an alignment position is given if 19 or more of the 39 residues are conserved according to the groups (ST), (PG), (KR), (QNED), (MC), (HFYW) and (AIVL)

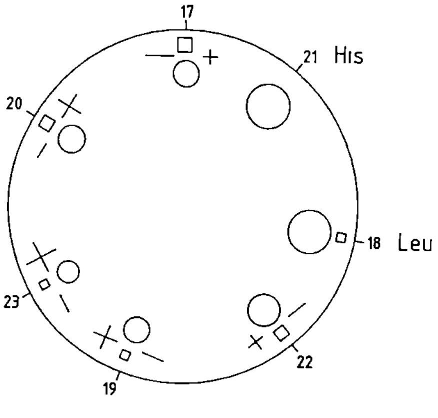


Fig 1 A 'helical' wheel representation of the predicted α-helical region in the 'finger' sequences The wheel is numbered according to the alignment in Table 1 The positions of the strongly conserved Leu 18 and His 21 are shown The symbols used to illustrate the amphipathic nature of the α-helix are (○) for hydrophobic residues (IALVMCHYWF), (+) for basic amino acids (KR), (−) for acidic/polar residues (QEDN), and (□) for (PSTG) The size of the symbols are proportional to the amino acid composition at each alignment position

An α-helical wheel representation[9] of the span (Fig 1) shows an obvious hydrophobic/hydrophilic sidedness, typical of helices observed in many three-dimensional protein structures One of the conserved histidines suggested for zinc coordination and a strongly conserved leucine are located on the hydrophobic side of the predicted α-helix This span may act as the principal finger region for interaction with nucleic acid

Two preliminary spectroscopic observations in this laboratory support the helical suggestion Circular dichroism curves of the TFIIIA protein as analysed by the method of Provencher[10] indicate the presence of 10–15 per cent α-helix A fluorescence quenching analysis of the 7S particle shows an approximate 30 per cent increase in intensity upon digestion of the RNA with ribonucleases or its salt-induced dissociation We interpret this result as evidence for changes in the environments of the two tryptophan residues (W28 and W177) in TFIIIA which are just at the amino-terminal side of the predicted α-helix Apparently their degree of exposure and structural flexibility is altered considerably upon binding to nucleic acid

RAYMOND S BROWN
PATRICK ARGOS

European Molecular Biology Laboratory,
Meyerhofstrasse 1,
D-6900 Heidelberg, FRG

1 Enver, T *Nature* 317, 385 – 386 (1985)
2 Berg, J M *Nature* 319, 264 – 265 (1986)
3 Harrison, S C *Nature* 322, 597 – 598 (1986)
4 Vincent, A *Nucleic Acid Res* 14, 4385 – 4391 (1986)
5 Brown, R S, Sander, C & Argos, P *FEBS LETT* 186, 271 – 274 (1985)
6 Miller, J, McLachlan, A D & Klug, A *EMBO J* 4, 1609 – 1614 (1985)
7 Rhodes, D & Klug, A *Cell* 46, 123 – 132 (1986)
8 Argos, P *EMBO J* 4, 1351 – 1355 (1985)
9 Schiffer, M & Edmundson, A B *Biophys J* 7, 121 – 135 (1967)
10 Provencher, S W & Glöckner, J *Biochemistry* 20, 33 – 37 (1981)

Data in graphs and tables

SIR—After reading Paolini's discussion on enzyme units and how to express them[1], my attention was drawn to a similar ambiguity in the literature, concerning the units in which data are expressed when shown in tables or plotted in graphical form

Basic mathematics tells us that the value of a physical quantity is equal to the product of a numerical value and a unit

physical quantity = numerical value × unit

According to the convention followed by the Royal Society[2], the expression used to define the numerical values of a physical quantity plotted on a graph (or the one which is placed at the head of a column of numerical values of a physical quantity in a table) should be a pure number, and the axis on the graph (or the head of the table) should be labelled with physical quantity/unit — for example, radioactivity/(d p m $\times 10^3$) — thus, making the numbers in graphs and tables dimensionless If this convention is not followed, that is, if the numerical values in graphs and tables are considered to be dimensional, then the axis on graphs and the heads of tables should be labelled with the appropriate units (d p m $\times 10^3$, and not d p m $\times 10^{-3}$ as is often the case)

Figure 1 summarizes several ways of expressing data in graphical form In this figure, molar was the chosen unit because conventional prefixes like nano, micro, milli, and so on, are very commonly used, which is not the case with d p m (for which Bq is now the recommended unit)

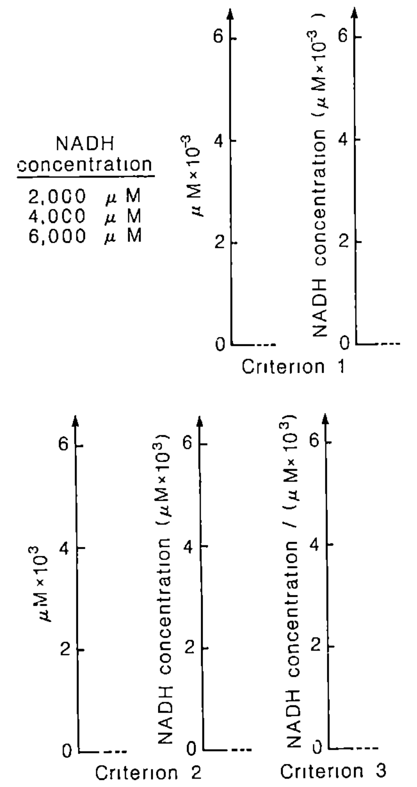


Fig. 1 Hypothetical experimental data showing common criteria for graphical presentation

and some other units Despite the method approved by the Royal Society (criterion 3) being mathematically the most precise and elegant, it is difficult to envisage its general use particularly outside the scientific community (in, for example, newspapers and television) In many (most?) biochemical papers, criterion 1 seems to be followed and even in well established books it is easy to find inconsistent usage here for example, in ref 3 criterion 1 is used in fig IV-29 but fig IX-2 uses criterion 2 Use of criterion 1 may be prone to ambiguous interpretation as shown in Fig 1 — with data in thousands of μM (indeed mM), we end up with the axis labelled μM × 10^{-3} when this criterion is followed (not in fact mM, but nM[1])

Although in virtually all cases it is easy to recognise which method has been used (and this may explain why several methods coexist), the indiscriminate use of different ways of labelling graphs and tables does not contribute to the clarity with which scientific knowledge must be presented In this respect, I would rule out criterion 1 because it is the most ambiguous

RICARDO FERREIRA
Department of Biochemistry,
School of Biological Sciences,
University of East Anglia,
Norwich NR4 7TJ, UK

1 Paolini M *Nature* 321, 568 (1986)
2 The Symbols Committee of the Royal Society *Quantities, Units and Symbols* (The Royal Society, London, 1971)
3 Dixon, M & Webb, E C *Enzymes* (Longman, London, 1979)

Silver from Chernobyl

SIR—I beg to differ with R H Flowers (*Nature* **323** 208, 1986) concerning the origin of ^{110m}Ag in fallout from Chernobyl

It is easily calculated that at most 400 grams of the fission product ^{109}Ag had been produced in the reactor core of Chernobyl-4 up to the moment of the accident On the other hand the reactor was equipped with about 200 in-core neutron detectors of the self-powered type with emitters composed of silver These emitters are up to 2 5 metres per detector and, in all, contain many kilograms of natural silver, about 48 per cent of which is ^{109}Ag

Therefore the ^{110m}Ag which has also been detected in significant amounts in the Netherlands, is indeed an activation product but almost entirely originates from molten detectors

H VAN DAM
Interuniversity Reactor Institute,
2629 JB Delft, The Netherlands

Chernobyl fallout in Debrecen, Hungary

SIR—Since 1952, total beta activity of fallout in precipitation has been measured in Debrecen without interruption[1,2] The precipitation is collected daily by a PVC funnel (40 cm diameter) and evaporated into standardized glass cups, where total beta activity is measured by an end window Geiger-Muller counter

On 29 April, 2 7 mm (340 cm^3) of rain was collected, but previous activity levels were unchanged However, the shower of rain on 1 May (0 48 mm) resulted in 1,450 Bq m^{-2} initial total beta activity (Fig 1) The next rainfall, on 9 May, was collected and processed in two parts (2 5 mm before noon and 3 1 mm afterwards) The initial activity of the first part was 2,500 Bq m^{-2} and that of the second part was only 160 Bq m^{-2}, which is impressive proof of the rapidity of change in tropospheric radioactivity and shows that the main mechanism of fallout deposition is rainfall The specific activities of these three samples were 2,645, 980 and 46 Bq l^{-1} respectively

We have identified the radioisotopes by the γ-ray spectra of the samples using a 100 cm^3 Ge (Li) detector surrounded by a 5 cm lead shield The isotopic composition of the fallout in the precipitations before noon on 1 May and 9 May proved to be the same, though the intensity ratios of short half-life radioisotopes were different

The residual activity on the later date was that of the isotopes ^{103}Ru, ^{95}Zr, ^{95}Nb, ^{106}Ru, ^{106}Rh and ^{134} ^{137}Cs

In all, after 10 May the total beta activity level was less than in the first three samples, and did not significantly exceed, those of past years after 25 May

E CSONGOR
A Z KISS
B M NYAKÓ
E SOMORJAI
Institute of Nuclear Research,
Hungarian Academy of Sciences,
H-4001 Debrecen,
POB 51, Hungary

1 Szalay, A & Czongor, E *Acta phys hung* 29, 407–413 (1970)
2 Csongor, E *Fizikai Szemle* 39, 11–20 (1985)

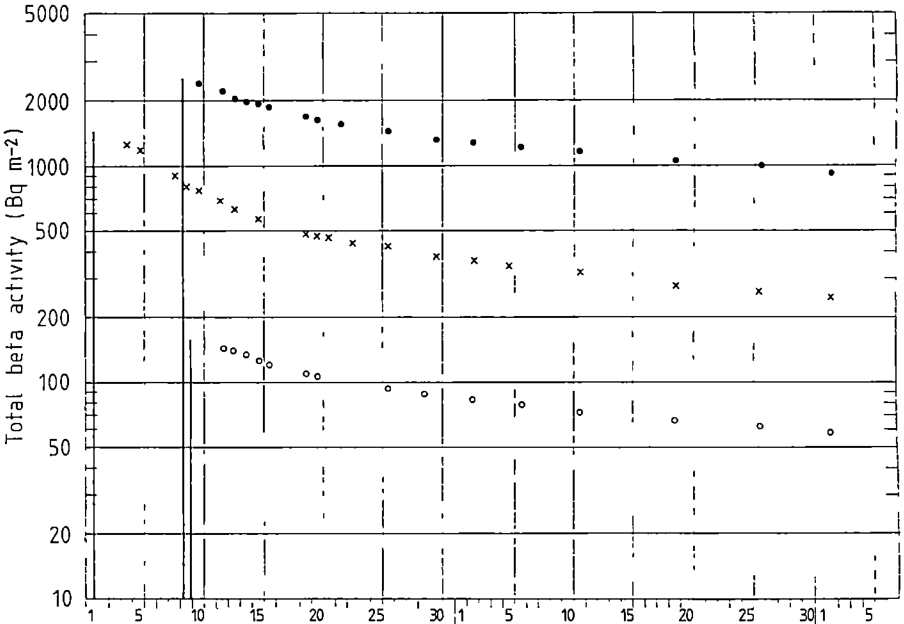


Fig. 1 Total beta activity of the Chernobyl fallout in the precipitation in Debrecen, Hungary during May–July 1986

Birth of a separate science

Peter Harman

Intellectual Mastery of Nature: Theoretical Physics from Ohm to Einstein. Vol. 1 The Torch of Mathematics 1800–1870; Vol. 2 The Now Mighty Theoretical Physics 1870–1925.
By Christa Jungnickel and Russell McCormmach. *University of Chicago Press: 1986. Vol. 1 pp. 350, $55, £46.75; Vol. 2 pp. 434, $65, £55.25.*

IN AN autobiographical sketch, written in 1891 towards the end of his exemplary career, Hermann von Helmholtz explained that he had sought "intellectual mastery of nature" in pursuing a career in science. Originally working principally in physiology, in physiological optics and acoustics, he had also made substantial contributions to physics. Helmholtz's classic memoir *On the Conservation of Force* (1847) provided a mathematical formulation of the principle of the conservation of energy, one of the grand unifying principles of nineteenth-century physics; and papers on hydrodynamics and electrodynamics confirmed his reputation as one of the greatest mathematical physicists of the day. Relinquishing physiology, he moved to Berlin in 1870 as professor of physics; he found the subject to be "the true basis of all proper science".

It was in theoretical physics, above all, that German scientists found "science as a vocation", and it was through the "torch of mathematics" (in Ohm's telling phrase) that they sought to realize their intellectual mastery of nature. At the beginning of the nineteenth century, German universities provided no more than elementary physics lectures for students who would go on to professional studies in medicine, law or theology. The main argument for educational reform was the hoped-for regeneration of the cultural and political reputation of the German states; its tool was the incorporation of research into the universities. The effect was the creation of scientific institutes and increasing specialization within the natural sciences. In their superbly written and exhaustively researched study, Jungnickel and McCormmach find the division of physics into theoretical and experimental specialities to be the distinctive feature of German physics in the nineteenth century.

Drawing on a rich supply of archival materials from universities and state bureaucracies, together with a thorough sampling of the physicists' scientific work, Jungnickel and McCormmach present a comprehensive and convincing picture of physics in nineteenth-century Germany. Some very complex science is described with economy and precision, but without undue technicality, and the authors have written an integrated narrative in which the scientific work of the physicists is portrayed in its institutional setting. The collaboration has proved remarkably successful, with the authors' interests (Jungnickel's more in institutional history, McCormmach's more in intellectual history) being effectively blended. In his *Night Thoughts of a Classical Physicist* (Harvard University Press, 1982) McCormmach has already presented a glimpse of the materials here meticulously documented, in the form of a fictional reconstruction of the life of a German physicist active from the latter part of the nineteenth century.

BBC Hulton Picture Library

Hermann von Helmholtz, "classical master".

Physical theory, or "natural philosophy", had long been a subject of study, yet only by the beginning of the nineteenth century was the term "physics" being used in the modern sense to denote a study of mechanics, optics and electricity which employed mathematical and experimental methodology. At that time physicists strongly emphasized quantification and the search for mathematical laws, and sought a unified science of physics based on the concept of energy and the programme of mechanical explanation. In seeking to give an integrated account of the scientific work in its institutional setting, Jungnickel and McCormmach concentrate on the physicists' work as an organizing principle. Research and teaching were carried out in institutes within the universities, and the book is structured around the development and proliferation of these institutes.

The pattern of development was remarkable. Early in the nineteenth century German physics was struggling to distance itself from the prevalent "nature philosophy". University professors and authors of texts tried to combat Hegel's contempt for the natural sciences. Around 1830 physics began to become established as an academic discipline, especially at Göttingen and the new university of Berlin. Gauss's collaboration with Wilhelm Weber in Earth-magnetic observations, Ohm's mathematical theory of electric currents and Franz Neumann's work on optics and electricity all contributed to establishing a coherent corpus of German physics. Much of the inspiration for mathematical physics came from France, the work of Laplace, Poisson, Fresnel and Fourier providing notable exemplars, but by the 1840s a German tradition was emerging. Helmholtz's unification of physics in terms of the conservation of energy drew on work by Gauss, Weber, Ohm and Neumann. German work on electricity was especially notable; in the 1840s, Weber, Kirchhoff and Neumann published substantial memoirs. As a tyro electrician in 1854, James Clerk Maxwell felt the need to study these works, the "views of heavy German writers". Weber's electrodynamics, Helmholtz's energy physics, Clausius's work on thermodynamics and the "mechanical" theory of heat, all strove for unifying concepts by which the foundations of physics could be established.

The subject prospered even though physics institutes suffered from shortage of space and resources, and after 1870 the development of theoretical physics as a study distinct from experimental physics owed much to the new teaching needs which faced physics institutes. The requirement for a second teaching post within individual universities led to the creation of "extraordinary" professors in theoretical physics. These positions were planned as transitional appointments for young scientists who were expected to move on to an "ordinary" professorship in experimental physics and to the control of a physics institute. But eventually, following the example of Berlin, institutes of theoretical physics came to be established with theoretical physics being understood as that part of the subject concerned with the precise measurement of physical constants as well as with mathematical physics. Specialization was considered acceptable if the twin sides of the subject were kept in balance: as Helmholtz urged,

> without mathematical physics, experimental physics is a very narrowly bounded science and affords little insight into the course of physical

> phenomena, while . . . mathematical physics without experimental physics, would likewise be a rather lame and unfruitful science.

This attitude shaped Helmholtz's evaluations of candidates for university positions. Jungnickel and McCormmach consider Max Planck as the first true specialist in theoretical physics (in the sense that his career developed entirely in the subject). Yet Planck was considered by Helmholtz to be an "excellent acquisition" in replacing Kirchhoff as Berlin's theoretical physicist in 1889; Planck was a "splendid man" because his work was closely based on the experiments of others, on the interpretation and development of experimental measurements.

The main intellectual influence shaping German physics after 1870 was the work of James Clerk Maxwell. Helmholtz's influential treatment of electrodynamics and evaluation of the unique importance of Maxwell's electromagnetic theory of light encouraged the research of Heinrich Hertz, first under Helmholtz's direction at Berlin and then at Karlsruhe, which ultimately led to his decisive experimental test in favour of Maxwell's theory in 1888. As a visitor to Helmholtz's institute anxious to gain experience of experimental work, Ludwig Boltzmann undertook experiments (on dielectric constants) which bore on the confirmation of Maxwell's theory. In his research on the kinetic theory of gases and the entropy concept, Boltzmann extended Maxwell's law of the distribution of velocities among the molecules of gas, creating statistical thermodynamics. And in Leiden, Lorentz developed the "electron" theory which blended elements from German electrodynamics with Maxwell's theory of the electromagnetic field. The theoretical initiatives of Lorentz and Boltzmann, both much sought-after in German institutes of theoretical physics (Lorentz remaining in Leiden, but Boltzmann seeking to assuage his depression by moving to and fro between Graz, Vienna, Leipzig and Munich), encouraged German physicists to look to Maxwell's electromagnetic theory and molecular mechanics for their inspiration.

By the end of the century German physics, especially theoretical physics, manifested classical coherence and also innovatory thrust. The lectures of the classical masters Helmholtz and Kirchhoff summarized and codified the achievements of the past. The endeavours of Planck and Wien, debating the rationale of energetic and electromagnetic foundations for physics, opened the door to the new physics of quanta and relativity of the twentieth century, itself a creation of German theoretical physicists yet shaped by the problems which were perceived to flow from Maxwell's work. Planck came to exercise a formative influence, succeeding Kirchhoff in Berlin, and on Helmholtz's death becoming the advisor for theoretical physics for the *Annalen der Physik*. His research on blackbody radiation led him to introduce the quantum of energy, although he remained cautious about its implications. Nevertheless he was an early and influential supporter of Einstein's special theory of relativity.

Jungnickel and McCormmach conclude their account with an overview of the developments which arguably constitute the high point of German physics, the creation of the general relativity and quantum theories in the first quarter of the twentieth century. The authors emphasize the confluence of intellectual and institutional developments. Einstein came to the Prussian Academy and to Berlin in 1913 as the outstanding theoretical specialist, and soon delivered the "scientific eggs" in his general theory of relativity. The emerging generation of theoretical physicists, including Pauli and Heisenberg, received their education at the hands of men who had pursued careers as specialist mathematical physicists, Sommerfeld at Munich and Born at Göttingen. If classical physics had achieved its apotheosis in the lectures of Helmholtz and Kirchhoff, the new generation could build creatively on the shoulders of giants. The "now mighty theoretical physics", as Wilhelm Wien described it in 1915, had attained status as a "separate science". ☐

Peter Harman is currently an Associate of Clare Hall, Cambridge, CB3 9AL, UK. He is author of Energy, Force, and Matter: The Conceptual Development of Nineteenth-century Physics *(Cambridge University Press, 1982).*

Naked appreciation

P.W. Hawkes

Principles of Charged Particle Acceleration. By Stanley Humphries Jr. *Wiley: 1986. Pp.573. $65.95, £63.15.*

THE difference between electron optics and the optics of charged particle acceleration is not unlike that between nudity and nakedness. It is possible to write a book on electron optics that sets out from a few premises and provides an elegant abstract description of the behaviour of electrons in the lenses, accelerating structures and analysers of which many instruments are composed. The typical example is P.A. Sturrock's *Static and Dynamic Electron Optics* (Cambridge University Press, 1955), in which it is not too fanciful to recognize the austerity of the classical nude.

In the optics of charged particle acceleration, on the other hand, the blemishes and homely imperfections of mere nakedness cannot be overlooked: many different kinds of particle may be present in the beam, the current density may be so high that the particles jostle one another in a way that renders the fundamental equations unmanageable, instabilities may set in and, in the great accelerators of recent years, even the irregularities of the underlying terrain cannot be neglected.

Stanley Humphries is well aware of the complexity of his subject matter and steers a middle course that accelerator physicists and engineers will be able to follow easily. Like Sir Thomas Browne, he believes that "where there is an obscurity too deep for our reason, 'tis good to sit down with a description, periphrasis or adumbration". He opens with concise accounts of particle dynamics in electric and magnetic fields, field calculation and materials properties. The next three chapters are concerned with lenses: round lenses, quadrupoles and sectors. Their effects on particle beams are first examined in terms of the elementary lens properties, after which the notions of acceptance and matching are introduced. The last of these chapters explains the transfer matrix formalism. Humphries does not go far into these optical matters, but this merely emphasizes that he has written a book on acceleration and has included only the necessary minimum about them.

Apart from a single chapter on the difficult theme of phase dynamics, the rest of the book offers a detailed and knowledgeable account of the various types of accelerator. Electrostatic accelerators and pulsed high-voltage, linear-induction accelerators, betatrons, radiofrequency accelerators and, finally, cyclotrons and synchrotrons: each of these is accorded a separate chapter, as are the resonant cavities and waveguides employed. Not only are the modes of action of the various machines clearly described, but the book is also a rich hoard of down-to-earth engineering detail that will make it used and appreciated far beyond the introductory course on the physics of charged particle acceleration that inspired it.

Humphries has brought together and successfully organized a mass of material in the first large single text to be devoted to his subject since *Linear Accelerators*, a multi-author work edited by P.M. Lapostolle and A. Septier and published by North-Holland in 1970. I suspect that many readers will need to go elsewhere to delve into the subjects of the earlier chapters in more detail, but they will surely forgive the author for preferring to save his space for the accelerators themselves. Altogether, this is an impressive and consummately useful volume. ☐

P.W. Hawkes is Directeur de Recherche at the Laboratoire d'Optique Electronique du CNRS, 29 rue Jeanne Marvig, 31055 Toulouse Cedex, France.

How mistakes are made

Leslie E. Orgel

Accuracy in Molecular Processes: Its Control and Relevance to Living Systems. Edited by T. B. L. Kirkwood, R. F. Rosenberger and D. J. Galas. *Chapman & Hall: 1986. Pp. 398. £48.50, $110.*

The frequency of errors in the transmission of DNA sequence information from one generation to the next can be as low as one per 10^8–10^{10} bases copied. Transcription and translation have substantially higher error rates, but are still far more chemospecific than comparable non-enzymatic processes. The measurement of error rates, the study of the enzymatic mechanisms that maintain them at low levels and the investigation of the cumulative effect of errors on the health of the organisms that harbour them have become the subject matter of a subspeciality of biochemistry. This subspeciality is the subject of the present multi-author volume. The "accuracy" of methylation, restriction, splicing and other post-synthetic modifications of proteins and nucleic acids is not discussed.

The empirical data on the accuracy of DNA replication, and the classic mechanism of error correction by exonuclease activity during replication are discussed in a chapter by Goodman and Branscomb. The role of DNA repair in maintaining the integrity of the genetic material is covered by Sedgwick. The accuracy of RNA synthesis is discussed by Anderson and Menninger.

Protein synthesis is dealt with in three chapters. Fersht covers the charging of tRNA and deals in detail with editing pathways that remove "incorrect" amino acids hydrolytically from charged tRNAs. Buckingham and Grosjean review the literature on the accuracy of mRNA–tRNA recognition and show that recognition depends on many factors beyond the complementary pairing between codon and anticodon triplets. Kurland and Gallant deal with the role of the ribosome in controlling the accuracy of protein synthesis.

In addition to these empirically based reviews, there are other, mainly theoretical contributions on such topics as the kinetic cost of accuracy (Ehrenberg, Kurland and Blomberg), selection for optimal accuracy (Kirkwood and Holliday) and the effects of errors on the integrity of genetic information transfer (Rosenberger and Kirkwood). A chapter on probabilistic thinking about enzyme kinetics by J. Ninio provides a number of qualitative insights that are not readily derived from the differential equations of traditional chemical kinetics.

This volume is a valuable addition to the literature of biochemistry because it brings together reviews of all of the important aspects of the "accuracy" problem. It is not always easy going — the introductory material is highly compressed and requires very careful reading, as do many of the more theoretical sections. But the determined reader will be rewarded with an excellent overview of an important but little-known area of biology. □

Leslie E. Orgel is Director of the Chemical Evolution Laboratory, The Salk Institute for Biological Studies, San Diego, California 92138, USA.

Solar systems

A.K. Turner

Physics, Technology and Use of Photovoltaics. By R.J. Van Overstraeten and R.P. Mertens. *Adam Hilger: 1986. Pp. 278. £35, $70.*

The potential of photovoltaic solar-energy conversion as a renewable energy source is widely recognized. Photovoltaic power is already finding remote-site application in telecommunications, navigation aids and water pumping, for example, predominantly using crystalline silicon-cell technology. Calculators and watches with amorphous silicon and cadmium telluride solar cells are also now common. Surprisingly, there are few good general texts on the subject, which makes Van Overstraeten and Mertens's contribution all the more valuable.

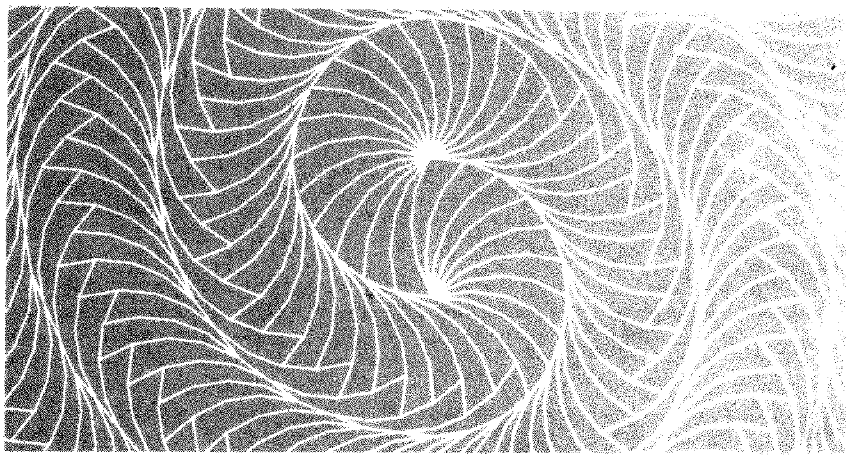
The book begins with a comprehensive account of the background to the physics of solar cells, made particularly interesting by the inclusion of a thorough description of the mechanisms which can account for loss of solar-cell efficiency. The theory is also developed separately for monocrystalline and polycrystalline silicon, amorphous silicon, gallium arsenide and other thin-film photovoltaic cells.

The method of manufacturing crystalline silicon photovoltaic modules is covered in great detail, from silicon purification and crystal growth, through cell fabrication, to the assembly of finished modules. In contrast, the section on thin-film technology is rather brief, with only a limited discussion of the various options. On the practical side, there are sections on measurement of the efficiency of light conversion of cells and modules, and their ability to withstand different environmental conditions, on cell interconnection in modules and on important factors in the design of photovoltaic systems for particular applications.

The book concludes with a summary of existing applications of photovoltaic systems which serves well to demonstrate the current dominance of crystalline silicon. This may well justify the authors' rather scant treatment of the emerging, potentially cheap, thin-film technologies such as amorphous silicon, cadmium telluride and copper indium diselenide cells which have yet to find large-scale commercial use beyond consumer applications.

Well written and logically structured, the book is easy to follow, and provides both a good text for the student and a useful reference work for the researcher. It describes the physics, technology and uses of photovoltaics for an audience anxiously awaiting a breakthrough in the availability of a low-cost, renewable energy source — we must hope that we don't have to wait too long. □

A.K. Turner is in the New Technology Division, British Petroleum Research Centre, Sunbury-on-Thames, Middlesex TW16 7LN, UK.



Singular construction — monohedral tiling (in which the tiles are all of the same shape) arranged into the semblance of a spiral; the prototile also admits many other, similar arrangements. The picture is reproduced from Tilings and Patterns, *an unusual and happy marriage of mathematics and aesthetics written by Branko Grünbaum and G. C. Shephard and published by W. H. Freeman. The book will be reviewed in a future issue of* Nature.

The many stages of carcinogenesis

D.G. Harnden

Introduction to the Cellular and Molecular Biology of Cancer. Edited by L.M. Franks and N. Teich. *Oxford University Press: 1986. Pp.458. Hbk £30, $59; pbk £15, $32.50.*

WILL Shakespeare started it: "One man in his time plays many parts". Since then analogies have run rampant. Now Beverly Griffin sees the seven ages of DNA — she does it well, and optimistically forecasts that in the seventh age the intricacies of the normal cell will be unravelled and lead to an understanding of cancer.

Perhaps I should try it — the seven ages of making a cancer textbook. First, the student, perhaps not "mewling and puking", but certainly confused by a subject such as oncology, where the range of disciplines is wide and information has to be picked up from many teachers and often from a variety of courses. Next, the teacher dedicated to a coordinated oncology course. Anyone who has tried, knows the problems. How can you even begin on DNA manipulation unless there is a firm grounding in DNA structure? Yet if you do not take much as read, there is not time to get from DNA to cancer therapy. And then, the brave editor, suggesting that the course should be converted into a book — the students are proferred more detail and time can be devoted to discussions. So distinguished experts send in their contributions, which somehow must be woven into an accurate and readable whole.

Much to their credit, Franks and Teich have succeeded to a large degree where others have totally failed. Their book will be of great value, not only to students but to anyone who wants a broad view of current knowledge of the scientific basis of cancer. The authors are indeed distinguished, the coverage is excellent and the prose very readable.

The groundwork is well laid by Franks himself, while chapters on metastasis and epidemiology ensure that the behaviour of cancer and the scale of the problem are well understood. But the heart of the book deals with carcinogenesis, cellular and molecular processes, and the nature of the cancer cell. The theme of stages set by Griffin recurs throughout the book. Multistage carcinogenesis is universally agreed, but each author sets these stages rather differently. For example, Adam sees radiation carcinogenesis in stages from the physical (10^{-12} to 10^{-18} seconds) to the cell and tissue stage which lasts months or years. Even growth factors operate in stages, and Waterfield spells them out from the external signal to the nuclear response. Though the details of clinical cancer are beyond the scope of the work, the link to clinical problems is handled briefly but clearly.

This book is therefore at about stage four in its evolution; very good but not yet excellent. The authors, not too "jealous in honour", invite comment for its future improvement. Perhaps they could look to the order of the chapters. For example, why does Greaves's excellent chapter on the biology of leukaemia come so early? It makes no concessions to the uninformed and confronts complex ideas about cell lineages, maturation arrest and expression of cell-surface molecules. The editors should also consider the difficult problem of how much background to fill in; some of the contributors expect sophistication of their readers, while others go back to first principles.

So onward to stage five. After that, stage six will be superb, "full of wise saws and modern instances". But what of stage seven? Let me predict — sans page, sans printed word, sans everything, except perhaps a computer terminal. □

D.G. Harnden is Director of the Paterson Institute for Cancer Research, Christie Hospital, Wilmslow Road, Manchester M20 9BX, UK.

Death on the Nile

D. M. Dixon

Science in Egyptology: Proceedings of the 'Science in Egyptology' Symposia. Edited by A. Rosalie David. *Manchester University Press/Longwood Press: 1986. Pp.525 £45, $77.50.*

IN THE course of excavations in Egypt and Nubia, carried out over several decades about the turn of the century, thousands of mummified and naturally preserved human bodies were recovered, ranging in date from pre-Dynastic times to the Christian period. Unfortunately, because of lack of specialist personnel on the staff of expeditions, the opportunities presented by this material were not fully exploited. With a few notable exceptions, excavators often simply selected for retention and later examination at home those bones which appeared to them to exhibit some abnormality or other point of interest. In general, the remains that did find their way to museums were for the most part neglected for years, during which specimens deteriorated and labels became detached.

Co-operation on any scale between Egyptology, science and medicine was slow in coming, and it was not until the mid-1960s that practical steps were taken to remedy the situation. The appearance of *Diseases in Antiquity*, edited by D. Brothwell and A.T. Sandison, and published in 1967 by Charles C. Thomas, marked a turning-point, although not all of the papers were concerned with Egypt. Shortly afterwards the work of Chiarelli and his colleagues at Turin, and the foundation of the Paleopathology Association in 1973 through the efforts of Aidan and Eve Cockburn, gave added impetus to interdisciplinary research on life in the ancient Nile Valley.

The Manchester Symposia of 1979 and 1984 have continued and expanded that work, the scope of the latter gathering being widened to include the discussion of techniques applied to pottery and other antiquities. This volume contains a printed record of all the papers presented in 1984, together with a selection of those from 1979. Inevitably, the contributions vary considerably in value, but there is much of interest here. Among the most promising developments is the extraction and cloning in bacteria of DNA from mummified tissue, discussed by S. Pääbo, and the possibilities this offers of determining, among other questions, the genetic relationships between individuals and populations. This is clearly a matter of some importance to scholars grappling, for example, with the complexities of the history of Egypt during the latter part of the Eighteenth Dynasty.

An obstacle in the way of examining mummies has hitherto been the understandable reluctance of museum curators to allow the use of techniques which involve the total or partial destruction of specimens in their care. The application of endoscopy, described by E. Tapp and his colleagues, has changed things almost overnight. It has also permitted the taking of biopsies from such areas as the lungs and the confirmation of diagnoses based on radiological examination.

While the range of techniques discussed in this book is impressive, and augurs well for research on the lives of ancient peoples generally, a note of caution is necessary. There is still a long way to go before Egyptologists can turn to their medical and scientific colleagues with the reasonable expectation of receiving satisfactory answers to their problems. In this regard Renate Germer's closing paper, "Problems of Science in Egyptology" makes salutary reading. As she reminds us, much of the information obtained through the use of methods discussed at the Symposia is at present of limited help to Egyptologists because "a great deal of basic research has yet to be done to determine the applicability of the scientific methods themselves". □

D. M. Dixon is a Lecturer in the Department of Egyptology, University College London, Gower Street, London WC1E 6BT, UK.

Glittering prizes for research support

David F Horrobin

Could public support for research be cheapened, and made more productive, by following the eighteenth-century precedent of the British government's prize for a means of measuring longitude?

THE crisis in British universities, the shortages of research funds and the insistent demands that science should serve the economic needs of the nation which supports it, ought to require a fundamental rethinking of our approach to the financing of science I see much evidence of panicky tinkering by both the scientific community and the government, but few indications of radical thought

We should perhaps consider what may be the first example of parliamentary funding of research, which is also one of the most successful Because it is such an excellent precedent, those in power might take it seriously The hero is Mr John Harrison, born in 1693 the son of a poor Yorkshire carpenter, a self-educated clockmaker and the toast of Neil Armstrong at a twentieth century 10 Downing Street dinner

Petition

In a petition to the British Parliament on 25 March 1714, Captains of Her Majesty's Ships, Merchants of London and Commanders of Merchant Men demanded a solution to the problem of measuring longitude at sea Failure was causing innumerable navigational disasters with the loss of ships, men, goods and battles The government, wisely, chose not to throw money at the problem by supporting research by experts who, by definition, had failed to solve it Instead, they arrived at an astonishing proposal a prize of £10,000 was offered for a method that would reliably measure longitude to within 1°, £15,000 for a method accurate to within 40′, and £20,000 for a method accurate to within 0 5° In 1714 those sums were truly astronomical Anyone winning the prize would become not merely comfortable but genuinely rich

The result was an outburst of private research on longitude The problem and the prize were known to everyone They are mentioned in Swift's *Gulliver's Travels* and a Hogarth cartoon shows a lunatic obsessed with them Theoretically, it was known that one way to solve the problem was to make a timepiece so accurate that it would tell the time at Greenwich even after many weeks at sea, much buffeting and drastic changes in temperature Comparison of that time with local time would then allow an accurate calculation of longitude

None of the experts believed that a timepiece could be so accurate An array of bizarre and complex proposals was put forward by the Fellows of the Royal Society and their friends, but all failed Only poor, provincial, self-taught John Harrison believed that a clock could be made to the required standards His story is well known his production of increasingly reliable, robust and accurate chronometers, their objective success in Royal Navy trials, his denigration by the eminent gentlemen of the Royal Society, the attempts by the Society's experts to deny him the prize legitimately won, the anger of the King when he heard about these academic machinations and Harrison's final triumph with grant of the full £20,000 at the age of 80 But the lessons to be drawn from this remarkable story are as valid now as they were in the eighteenth century

- The announcement of a cash award sufficiently large to make the winner enormously wealthy will generate a phenomenal outpouring of research directed to the solution of practical problems Such research will be privately financed and will cost the donor nothing but the cost of the award
- Such an award will lead to unprecedented cross-fertilization of fields of endeavour If the prize is large enough, scientists whose speciality may appear remote from the problem will be stimulated to think about it
- The definition of success must be crystal clear, practical and open to verification by non-experts The solution need not be brilliant or sophisticated, nor need it meet with the intellectual approval of experts The only condition is that it must work
- There must be no limitation on the categories of people allowed to succeed The challenge must be open to non-experts and to the sons of poor carpenters

Prizes

The government should decide what problems it wants solving In my field of medicine, obvious problems whose solutions would save a great deal of money include schizophrenia, eczema, multiple sclerosis and Alzheimer's disease People from other fields should be able to produce long and comparable lists Economists could then work out what each particular problem costs the nation, and a prize could be offered for a practical approach which would either eliminate the problem or reduce the cost of solution The prize should be the value of the savings made during one full year

Some of the prizes might turn out to be very large indeed, tens or even hundreds of millions of pounds Even greater amounts would be possible if, say, the European Community, the United States and other countries acted together, as they might well do if the scheme cost little or nothing The prize should be given tax free to the individual, company or group producing an answer

The advantages of the scheme would be many, but include

- The cost of initiating the scheme would be trivial After the first year, a successful solution would save millions
- Researchers of all kinds would be encouraged but not forced to work on practical problems They would also be encouraged to extend their range of expertise
- There would be large inflows of private venture capital into research as individuals, companies and others chose to support research directed towards the solution of problems attracting the larger prizes
- There would probably be rapid progress in basic science Many now suffer from the illusion that basic discoveries always precede practical research Historical evidence suggests that the traffic is just as frequently in the other direction We have forgotten the story of Pasteur Brilliant people devoting themselves to practical problems, such as the spoiling of wine or dying silkworms, may end up by discovering new fundamental principles
- The enterprise would be cash-generating rather than cash-consuming Governments would be delighted with science instead of looking at it with a jaundiced eye
- Last, but by no means least, the huge cash prizes would put money into the hands of the most practically creative members of society The money would certainly be spent in interesting ways, on further research, on enterprising investment or even on innovative techniques of personal consumption

The idea would cost almost nothing to put into operation It would rejuvenate science and transform attitudes to science It should be given a try □

David F Horrobin is managing director of Efamol Ltd, Woodbridge Meadows, Guildford GU1 1BA, UK

Uncertainties in the smoke source term for 'nuclear winter' studies

Joyce E. Penner

Lawrence Livermore National Laboratory, University of California, Livermore, California 94550, USA

Climate models have shown that the effects on climate of a major nuclear exchange depend on the quantity and optical properties of the smoke that is dispersed into the global atmosphere. Published estimates for each of the factors necessary to determine the smoke optical depth yield a wide range of values, which allows the prediction of either comparatively minor effects on climate or massive effects.

THE potential effects on climate of large amounts of smoke injected into the atmosphere following a major nuclear exchange have been widely analysed[1-4] Although simplifications and uncertainties still exist in the application of climate models to calculate the effects of smoke, many of the simplifications that had to be made in the first studies have now been corrected These improved climate models have shown that the effects of smoke on climate depend on the quantity and optical properties of the smoke that is generated and dispersed into the global atmosphere

The amount of smoke and its optical properties can be summarized by the average optical depth that would result if the smoke were dispersed throughout half the Northern Hemisphere If the extinction optical depth $\bar{\tau}_e$ is taken as the sum of the scattering and absorption optical depths, it may be calculated from $\bar{\tau}_e = (k_s + k_a)S/A$, where k_s and k_a are the scattering and absorption cross sections of the smoke (in $m^2\,g^{-1}$), S is the amount of smoke (in g) and A is the area blocked by the smoke cloud (taken as half the area in the Northern Hemisphere, or $1\ 28 \times 10^{14}\ m^2$) The quantity of smoke may be calculated from $S = \varepsilon F(1 - f_r)$, where ε is the emission factor of smoke (g smoke per g of fuel burned), F is the amount of fuel burned, and f_r is the fraction of smoke removed by precipitation in the convection column above the fires and in the first few days after the war, during which the smoke is presumed to spread out to global scales The original 'baseline' analysis of Turco *et al*[2] resulted in an estimate for $\bar{\tau}_e$ of ~6 for smoke produced in urban fires Their smoke had a single scattering albedo of ~0 5, so that $\bar{\tau}_a = k_a S/A$ was ~3 0 Consideration of the additional smoke from wildlands fires, long-term fires and dust from surface bursts increased the estimate for $\bar{\tau}_e$ to ~8 (Turco *et al*[2] spread their smoke over the entire Northern Hemisphere, so that their published optical depths differ from these by a factor of 2) All subsequent analyses have similarly implied 'best estimates' for $\bar{\tau}_e$ that were substantially greater than 1 Here I review the estimates for the various factors which contribute to $\bar{\tau}_e$ and $\bar{\tau}_a$, to obtain reasonable bounds on the range of magnitudes consistent with current knowledge This range encompasses values that may be associated with major climatic effects if a large fraction of the available urban combustible load is burned On the other hand, within the bounds set by current analysis, comparatively minor effects are also possible, especially if the targeting of weapons avoids refineries and other large storage facilities that contain petroleum or other fossil fuels I then recommend several areas for research that could lead to more certain estimates of the effects of such a war

Amount of fuel burned in urban fires

As pointed out in ref 5, several methods have been adopted for estimating the amount of fuel that might burn in urban fires These methods are not mutually consistent In the method adopted by Turco *et al*[2], which yields the highest fuel estimate, the amount of fuel burned in urban fires is determined from the product $F = \mathrm{FL} \times f_b \times A_i \times \mathrm{SF}$ where FL is the average areal fuel load within the burned-out region (in $g\ m^{-2}$), f_b is the fraction of fuel that is consumed by fire within the area that burns, A_i is the area that is initially ignited by the fireball (in m^2) and SF is the average areal spread factor for the fires

An overestimate by this method may be caused by at least two factors Most often, no account is taken of the overlap of burned areas when detonations take place near one another Secondly, the entire ignited area and its radially expanded spread is assumed to coincide exactly with the urban fuel bed and with the average fuel load FL This assumption may be seriously in error, for example, for targets such as airports that generally reside on the outer edges of cities It is often argued that these effects are mitigated by the choice of a 'conservative' value for A_i—that is, one that corresponds to the area that would be ignited by a thermal fluence of $20\ cal\ cm^{-2}$, rather than the area associated with a thermal fluence of $7\text{-}10\ cal\ cm^{-2}$, which is considered sufficient to ignite at least the lighter fuel elements such as paper and twigs In Nagasaki, where the presence of hills restricted the fire ignition area[6], the actual area burned ($A_i \times \mathrm{SF}$) corresponded to the area which would have received a thermal fluence of $20\ cal\ cm^{-2}$, whereas in Hiroshima, the area burned corresponded to the area which received only $7\ cal\ cm^{-2}$ It would therefore seem that the area corresponding to $20\ cal\ cm^{-2}$ is indeed 'conservative', if $\mathrm{SF} = 1$

It is not known, however, whether the two effects mentioned above (overlap and improper average values for FL) would indeed be balanced by an underestimate for $A_i \times \mathrm{SF}$ Several lines of inquiry suggest that the overestimate may be significantly larger than the factor of 2 underestimate made by using a fluence area corresponding to 20 rather than $7\text{-}10\ cal\ cm^{-2}$ The analysis in ref 7, which consisted of a large attack on cities, suggests that consideration of overlap may reduce the value of A_i by as

Table 1 Average combustible fuel load in cities

Authors	Fuel load ($kg\ m^{-2}$)
Turco *et al*[2]	
Baseline case	33 5*
100-Mtonne, city centre	200 0
Crutzen *et al*[10]	40 0†
NRC[3]	40 0
Reitter *et al*[8]	
Detroit, centre	34 5
Detroit/San Jose, suburbs	10 2

* Average of $100\ kg\ m^{-2}$ in 'city centres' and $30\ kg\ m^{-2}$ in suburbs
† This estimate was not used in the final analysis of ref 10 (see text)

much as a factor of 4 Also, overlap of ignition areas may reduce fire spread, if intense firestorms are generated which create inflowing winds On the other hand, with more intense fires, the fuel may burn more thoroughly, increasing f_b

The value used for the average fuel load can be checked by an analysis of some of the fire spread modelling results of Reitter *et al*[8] (see also Appendix 3A of ref 4) All previous studies of the amount of fuel that might burn have assumed average values for FL of ~40 kg m^{-2} (see Table 1), although values from 10 to 400 kg m^{-2} are quoted as possible[3] Table 1 also shows the average areal fuel loads within the ignited areas corresponding to a 1-Mtonne nuclear explosion over the centre of Detroit and several 1- and 0 5-Mtonne bursts over detonation points above suburban Detroit and central San Jose, as taken from Reitter *et al*[8] These data imply that values closer to 10 kg m^{-2} should be used for most urban and suburban areas Furthermore, if the fuel loads for Detroit are correct, values for FL of ~40 kg m^{-2} are appropriate only for weapons directed at the centres of large cities The average fuel loads from Reitter *et al* consider only those areas occupied by buildings, so that these average values do not account for any decrease in FL due to targeting on the fringes of cities or near lakes or parks which would have lower average fuel loads The fuel loads were developed from surveys taken in the late 1960s, but recent analyses of fuel loads in San Jose[9] are similar to the average fuel load used by Reitter *et al* for that city The Detroit fuel loads used by Reitter *et al* seem surprisingly low, but result from averaging the fuel loads in the centre of that city (which range up to 160 kg m^{-2} over small areas) with values near 10 kg m^{-2} in outlying regions Note that Detroit has a population of >4 million There are only 39 urban centres in the world with a population of >3 million, and only 80 cities in the NATO and Warsaw Pact countries with populations of >1 million The oft-quoted "100-Mtonne central city" model of Turco *et al*[2], assumed values for FL of 200 kg m^{-2} in 100 cities these loads appear to be overestimated by a factor of 5 Of course, one must check whether European cities or cities in the eastern United States contain much higher fuel loads than those quoted here for Detroit But it seems highly probable that the estimates for FL used previously are too large, especially in view of the analysis of total fuel load outlined below Crutzen *et al*[10] also rejected this method in favour of the inventory approach outlined below Finally, note that the above analysis for FL assumed targets which were entirely contained within the urbanized areas occupied by buildings Consideration of actual target locations, some of which will occur on the fringes of cities and some of which will fall near lakes or other low-fuel-density areas, will further reduce the estimate for FL

Significant further reduction of the uncertainties using this approach requires a detailed analysis on a city-by-city basis, with consideration of specific target locations, fuel loads and overlap of fire areas Here we consider an alternative approach, in which total combustibles are estimated directly and then a fraction is assumed to be ignited and burned

The total inventory of combustibles has been estimated by Crutzen *et al*[10] and by Bing[5], using different methodologies Pittock *et al*[4] also estimated inventories, but as ref 4 is primarily a review of ref 10, they will not be separately quoted here Crutzen *et al*[10] used production figures for various raw materials and estimates of their lifetimes to obtain estimates for the total abundance of cellulosic materials, polymeric materials and asphalt Bing[5], on the other hand, gathered data from surveys of fuel loads in various types of structures and their contents for the United States, and extrapolated these data to Europe and the Soviet Union The two sets of published figures are not directly comparable because Crutzen *et al* estimate the amount of celluosic and polymeric materials in the developed world, whereas Bing's estimates refer only to the NATO and Warsaw Pact countries Crutzen *et al* and Bing also separately estimate the amount of petroleum available to burn, including petroleum stored as primary stocks and as secondary stocks The figures of Crutzen *et al* refer to the amount of petroleum stored globally, wheras Bing's numbers again refer only to that fraction contained within the NATO and Warsaw Pact countries In order to consider similar models, I have reduced the inventories published by Crutzen *et al* for the developed world by the ratio of the population of the NATO and Warsaw Pact countries to that of the developed world Their estimates for petroleum were similarly reduced by the ratio of consumption rates in NATO and Warsaw Pact countries to that of the world As shown in Table 2, Crutzen *et al*'s inventory still implies 2 5 times more cellulosic material than does Bing's Remarkably, as the methodologies were different, the estimates for petroleum-derived materials are similar The total amount of fuel in primary stocks of petroleum is fairly well known and the two estimates are comparable The amount of fuel in secondary stocks of petroleum is less well known Two different sources of data have led to a difference of a factor of 2-4 for fuel in this category, as shown in Table 2

The methodologies used in these two studies have obvious difficulties, and it is not clear which method is more appropriate Note, however, that the totals for cellulosic and polymeric materials assumed by Crutzen *et al*[10] are not entirely consistent with the fuel load estimates derived from Reitter *et al*[8] For example, we may use the average areal fuel load for urban and suburban areas from ref 8, together with the total urban area in cities with population greater than 2,500 in the United States, 135,000 km^2, to arrive at a combustible load for the United States of 1,350 Tg Consideration of 50 city centres with fuel loads similar to that of Detroit might increase this total to 2,000 Tg This number is close to the value derived by Bing[5] for the United States (2,119 Tg), and thus lends confidence to his estimates On the other hand, when Crutzen's numbers for the developed world are scaled by the ratio of population in the United States to that in the developed world (0 225) we arrive at 4,400 Tg, which is at least twice as large as the estimate above In the analysis below, both numbers will be used to estimate the range of optical depths that are possible, given current uncertainties

Table 2 Inventory of total available combustibles in NATO and Warsaw Pact countries (Tg)

Authors	Cellulosic materials	Primary petroleum stocks	Secondary petroleum stocks	Polymeric materials
Crutzen *et al*[10]	16,500*	462†	198-462†	574*
Bing[5]	6,444	480	100	753

* Reduced from the estimate for the 'developed world' in ref 10 by the ratio of populations, 0 87 (G Bing, personal communication)

† Reduced from the estimate for the whole world in refs 4 and 10 by the ratio of consumption rates in NATO and Warsaw Pact countries to that of the world, 0 66

Table 2 summarizes the inventories of combustibles in NATO and Warsaw Pact countries, derived by these two methodologies To allow consideration of the range of smoke absorption properties from various fuel types, Table 2 divides the inventories into cellulosic fuels, petroleum-derived fuels and liquid fossil fuels This last category has been subdivided into primary and secondary stocks of petroleum Secondary stocks are considered to be distributed with other fuels, whereas primary stocks of petroleum are considered separately in order to calculate the effect of a concerted effort to avoid or include these targets (see estimates below)

Typically, only a fraction of the total available combustible material might actually burn in active, flaming combustion Here I assume this fraction to be 25% of the distributed fuels (cellulosic and polymeric materials, and secondary stocks of petroleum) for both the high and low estimates In this way the range of optical depth derived can be considered to be independent of any particular model, although more (or less) fuel might burn if the warring nations made a concerted effort to try to ignite (or avoid burning) the available fuel A 25% fraction might come about, for example, by associating ~65% of the total fuel with people who live in cities (the average proportion of city dwellers for Europe, the Soviet Union and the United States), and then burning 80% of the total fuel in cities, half in active, flaming combustion and half in longer-term, smoldering combustion Alternatively, most city fuels might burn in active combustion but, because of clustering of targets and overlap of ignition areas, only 40% of the fuel in cities actually ignites and burns The optical depths derived below could easily increase if the war involved, for example, attacks on oil rigs or on oil-producing countries whose whose fuel stores are not part of the inventory considered here Similarly, if attacks included population centres outside the countries included here, the optical depths would increase However, the NATO and Warsaw Pact countries contain 85% of the population of all countries in the developed world, and as the present model burns 40-80% of the urban fuels in these countries, it is certainly large enough to be considered a 'major' nuclear war On the other hand, ~30% of the population in the develped world is concentrated in the 100 or so largest cities (see Table 3 1 in ref 4), so that if 100% of the fuel of these cities is burned in active, flaming combustion, nearly the same total optical depth could be produced However, the purpose here is not to develop war models, but to show the range in optical depth expected from current knowledge for a given model These ranges can easily be adapted to other models

Table 3 Smoke emission factor

Authors	Per cent of fuel
Turco *et al*[2]	2 7*
Crutzen *et al*[10]	
Wood	1 5
Oil, polymers, etc	7 0
NRC[3]	
Wood	3 0
Oil, polymers, etc	6 0
Range of values in flaming combustion[10]	
Wood	0 085-2 5
Oil	2-10
Plastics	1 2-50

* Weighted average of the 'net emission factors' of 1 1% for urban centres and 3 3% for suburbs These values may include some allowance for scavenging by rain Emission factors in an earlier version of ref 2 were 2 5% for city centres and 5% for suburbs

Smoke emission factor

The appropriate smoke emission factor in a large-area urban fire depends on a number of poorly estimated and poorly known factors Emission rates can vary[11,12], depending on the type of fuel, the ambient air temperature, the availability of oxygen, the radiant intensity (as determined by the proximity of nearby fires), the geometric arrangement of fuel, and so on Only very limited data from large fires are available, so most studies have used values consistent with the range of emission factors measured in laboratory-scale fires (see Table 3) These might be underestimates, if oxygen availability is truly limited in a large-area fire On the other hand, Carrier *et al*[13] have argued that oxygen availability should not be an issue, given the turbulent motions above the fire In view of the lack of credible data for smoke emission factors from large-area fires, the values adopted here are estimated from the limited available data, and are highly uncertain Table 3 also includes the range of estimates of the emission factor compiled in ref 10, primarily from laboratory data In the analysis below, I adopt a range of values for the emission factor, consistent with the best estimates chosen in refs 3 and 10, however, larger uncertainties apply because of the possible inapplicability of these emission factors to large-area fires

Table 4 Absorption and extinction coefficients for smoke from urban fires

Authors	k_a ($m^2 g^{-1}$)	k_e ($m^2 g^{-1}$)
Turco *et al*[2]	2 9	5 8
Crutzen *et al*[10]		
Wood	3 3	6 8
Oil, etc	7 0	10 5
NRC[3]	2 0	5 5
Penner and Porch[15]		
Wood, no coagulation	1 5	6 6
Oil, etc, no coagulation	5 6	9 5
Wood, after coagulation	1 3	4 0
Oil, etc, after coagulation	1 8	4 0

Optical properties of smoke

Just as the emission factor for smoke depends on the burning conditions and type of fuel, so does the chemical, morphological and optical character of the smoke Nevertheless, various authors have estimated the absorption and extinction coefficients for smoke, based on a variety of measurements that have tended to emphasize the data available from smoke emitted under flaming conditions The estimates are summarized in Table 4 In most recent evaluations, the optical properties of wood smoke are distinguished from those of smoke from fuels such as oil, plastics, and other polymers whose chemical structure has little available oxygen These latter fuels tend to produce much blacker smoke Table 4 shows wide variations in the estimates of the absorption and extinction coefficients for fresh smoke

In addition to variations in average optical absorption and extinction from different evaluations of these properties for fresh smoke, two mechanisms may act to make aged smoke less absorbing (see ref 4) The first mechanism is coagulation, which may act on short timescales (in very dense smoke plumes) or on longer timescales (from days to a week in the spreading global plume) to create larger particles These larger particles would absorb and scatter less radiation, if they are spherical Because some smoke particles are quite oily (and therefore spherical), while others appear as fluffy or chained agglomerates, it is not possible to predict the effects of coagulation on optical properties Chained agglomerate particles might become spherical if they coagulated with oily smoke particles, or if they were to condense and re-evaporate in a cloud—a process which might allow the chains to collapse[14] To the extent that the agglomerates remain in a chained formation, their absorption properties may not change significantly. Thus, in the following, I adopt two extremes In the first case, coagulation is assumed to have no effect on optical properties, in the second, coagulation is assumed to reduce extinction and absorption by the amounts estimated in ref 15 for several days of coagulation This additional consideration widens the discrepancy between the lowest and highest estimates of absorption coefficient by an additional factor of ~3 for the highly carbonaceous, absorbing smokes The extinction coefficients differ by a factor of >2 The absorption coefficient for less absorbing smoke is not significantly changed by coagulation

Fraction of smoke scavenged by rain

The last factor which contributes to estimates of the average optical depth is the amount of smoke that is removed by precipitation occurring in the smoke plume over the fire and in the

Table 5 Fraction of smoke removed by rain

Authors	Fraction removed
Turco *et al*[2]	
Suburban fires	0.25
Firestorms	0.50
Crutzen *et al*[10]	0.30
NRC[3]	0.50
Cotton[17]	~0.02*
Penner and Edwards[18]	0.50†

* Fraction of smoke particles that enter cloud water by phoretic scavenging. The amount removed by precipitation may be less than this.

† Fraction of smoke particles that enter cloud water by nucleation scavenging. The amount removed by precipitation may be less than this.

first few days after the war. Several authors have estimated that, especially for large, intense fires, large quantities of water will condense above the fire[16,17], but there have been few attempts to quantify how much smoke is removed. In both Hiroshima and Nagasaki, a 'black rain' fell, coincident with the fires which followed the nuclear blasts of August 1945[6]. The black rain is presumably smoke that has been scavenged by rain. It is not known how much of the smoke was scavenged or whether the experience in Nagasaki and Hiroshima should be considered to be typical.

The amount of smoke scavenged by rain depends, once again, on properties of the smoke which are poorly known. For example, the number of smoke particles that act as condensation nuclei for cloud drops depends on the highest level of supersaturation attained above the fire. This depends on the total number of smoke and debris particles, their size distribution, and their affinity for water. Furthermore, the highest level of supersaturation depends on the updraft velocity within the plume as well as the growth rate of the drops which form. As the size and composition of smoke particles from the large-scale urban fire are poorly known, it is difficult to predict how much nucleation scavenging might occur. Initial studies, which assume idealized spherical particles, indicate that for conditions considered typical of smoke above fires, all particles with radii greater than 0.1 μm (or ~50% of a log-normal distribution with a mode radius of 0.1 μm) might be scavenged in this manner[18]. However, further studies with a more realistic description of smoke particles are needed.

Once drops have formed, other mechanisms may also act to attach smoke particles to cloud drops. These include electrical capture, phoretic forces and turbulent motions, although Cotton[17] has estimated that <2% of the smoke particles might be captured phoretically. Cloud drops may or may not form precipitation-sized rain drops. The probability of this occurring depends on the initial size of debris and smoke particles and on the number that become nucleated to form drops. Pruppacher[19] estimated that, although most of the particles would enter cloud drops by nucleation scavenging, rainout of the smoke would be relatively small because of overseeding effects. However, only one size distribution was considered and the updraft velocity was only 1 m s^{-1}. Consideration of more realistic fire plume conditions may alter this conclusion. Once the drops become large enough to obtain a significant fall velocity, they may capture more smoke particles by impaction scavenging. The probability of this occurring also depends on the size of smoke particle (with larger particles being more likely to be scavenged).

The capture mechanisms described above apply to warm-rain precipitation only. Additional mechanisms and pathways for capture must be considered in the case of ice formation, which can also lead to release of the smoke particles. Because the saturation vapour pressure over ice is less than that over water, when ice begins to form, water on drops may evaporate and recondense on ice particles, thereby releasing any smoke that was previously captured by nucleation scavenging. Recently G. Tripoli (personal communication) has used the cloud model developed by Cotton[17] to estimate that this process could limit any rainout of the smoke above the fire to <20%. In addition to the near-immediate rainout of smoke above the fire, mesoscale atmospheric circulations may be set up which lead to rainout on longer timescales[20].

The theoretical analysis of scavenging and rainout is complex and difficult, and further work is needed. For this reason, many authors have simply guessed a fraction for smoke that might be removed by rainout. These guesses range from ~0 to 50% (see Table 5), although the real range of possibilities might include values up to 100% in some cases[21]. Here I consider the range from 0 to 50%. Obviously, the range of average optical depths obtained could be larger, for example, if rainout removed 90% of the smoke.

Calculated optical depths and climate

If we combine all the choices described above, emphasizing the smallest factors in one case and the largest factors in the second,

Table 6 Calculated range of average optical depth from urban fires

Category	k_a (m^2 g^{-1})	k_e (m^2 g^{-1})	ε (g g^{-1})	F (Tg)	$(1-f_r)$	$\bar{\tau}_a$	$\bar{\tau}_e$
Wood							
High	3.3*	6.8*	0.03†	16,500*/4	1.0‡	3.19	6.57
Low	1.3(1.5)§	4.0(6.6)§	0.015*	6,444‖/4	0.5†	0.12(0.14)	0.38(0.62)
Polymers, plastics, etc., and secondary stocks of petroleum							
High	7.0*	10.5*	0.07*	1,083¶/4	1.0‡	1.04	1.55
Low	1.8(5.6)§	4.0(9.5)§	0.06†	674#/4	0.5†	0.07(0.22)	0.16(0.38)
Primary stocks of petroleum							
High	7.0*	10.5*	0.07*	480‖	1.0‡	1.84	2.76
Low	1.8 (5.6)§	4.0 (9.5)§	0.06†	462*	0.5†	0.19 (0.61)	0.43 (1.03)

* From refs 4 and 10.

† From ref. 3.

‡ Based on ref. 17.

§ From ref. 15. Numbers in parentheses refer to the case with no coagulation.

‖ From ref. 5.

¶ This number is the sum of the average of high and low estimates for secondary stocks of petroleum from ref. 10 and Bing's[5] estimate for polymeric materials (see Table 2).

# This number is the sum of Bing's[5] estimate for secondary stocks of petroleum and Crutzen *et al*'s[10] estimate for polymeric materials (see Table 2).

Table 7 Average optical depth

Model	$\bar{\tau}_a$	$\bar{\tau}_e$
Distributed fuels only		
High	4 2*	8 1
Low	0 2†(0 4)‡	0 5(1 0)
With primary stocks of petroleum		
High	6 1§	10 9
Low	0 4‖(1 0)	1 0(2 0)

* Equivalent to 270 Tg of smoke with $k_a = 2\ \mathrm{m^2\ g^{-1}}$
† Equivalent to 12 Tg of smoke with $k_a = 2\ \mathrm{m^2\ g^{-1}}$
‡ Optical depths in parentheses refer to the case with no coagulation
§ Equivalent to 388 Tg of smoke with $k_a = 2\ \mathrm{m^2\ g^{-1}}$
‖ Equivalent to 24 Tg of smoke with $k_a = 2\ \mathrm{m^2\ g^{-1}}$

we obtain the range in absorption and extinction optical depths shown in Table 6 Table 6 lists separately the optical depths from cellulosic fuels, from distributed fuels producing highly carbonaceous smoke (polymeric materials and secondary stocks of petroleum) and from primary stocks of petroleum In the case of distributed fuels, one quarter of the total abundance in the NATO and Warsaw Pact countries, or 40–80% of the fuel in cities, is assumed to burn As shown in Table 6, the burning of polymers and petroleum may contribute significantly to the total optical depth To show its effect, we consider two models In the first, the contribution of primary stocks of petroleum to the total optical depth is not included This might result if the warring nations specifically tried to avoid targets, such as refineries, that would add disproportionately to the optical depth As warring nations have never yet refrained from attacking the economic base of their enemies, this may be an unlikely model, but it is interesting to calculate the consequences should studies of 'nuclear winter' cause nations to re-evaluate their war plans In the second model these targets are all included, so that 100% of the primary stocks of petroleum are burned Table 7 summarizes the high and low estimates of optical depth for these two cases If the primary stocks of petroleum are not included, the absorption optical depth varies from 0 2 to 4 2 Including these stocks increases the range of absorption optical depths to 0 4–6 1 In the first case, the low estimate is equivalent to 12 Tg of smoke with the optical properties assumed in ref 3 This increases to 24 Tg of smoke if primary stocks of petroleum are included This case is close to the lowest amount of smoke (20 Tg) considered by Malone *et al*[22] in an advanced three-dimensional climate simulation Their results are consistent with widespread temperature changes of −4 to −6 °C over the continents in summer The largest average optical depth calculated here is equivalent to almost 400 Tg of smoke (assuming the absorption coefficient from ref 3) This is somewhat less than the largest amount of smoke assumed by Malone *et al*[22] (500 Tg) and would, according to their results, lead to profound climate changes, particularly as its removal would be inhibited by changes to atmospheric stability

The range of values calculated here is disquieting, because I tried to choose values for each of the various factors that were thought to be a 'best estimate' by at least one of the authors cited here Although the range derived is similar to that quoted in ref 3, they developed their estimates from uncertainties about a median value, whereas the present estimates are developed using the best estimates from a number of authors The calculated range would be larger if we also considered the uncertainty estimates for each of the factors This is especially true for the range of values chosen for the emission factors In this case, the range of published values is not very large, but, because so few relevant data are available, the published range may not even include the correct value There are also uncertainties caused by the lack of good data on the optical properties of smoke and the effect of the clouds formed above large fires Furthermore, as discussed above, little is known about the scavenging and rainout of smoke in fire plumes Good data on the properties of smoke from large fires will only come from large fire experiments, but great care is needed in the design and interpretation of the large-scale fires which will be used in the re-analysis of the smoke source term The planned experimental programs sponsored by the United States Defense Nuclear Agency should be helpful but must not stop after only the first few experiments We must try to understand the more complex situations that will exist in a real nuclear fire In addtion, more and greater emphasis must be placed on understanding scavenging and rainout Here, progress may come through the development of advanced modelling capabilities, coupled with verification by large-scale fire experiments

Conclusions

Although our lowest estimates for $\bar{\tau}_a$ and $\bar{\tau}_e$ may produce only minor climatic effects, models can easily be constructed in which more fuel is burned, so that even in the low-estimate case, the estimate for $\bar{\tau}_a$ would correspond to a major climatic impact if the war takes place in the spring or summer On the other hand, it is entirely possible that such major effects could be avoided if the low estimates are correct and if targets such as refineries, oil and gas production fields and coal storage areas are avoided The effects on climate could also be lessened if the war took place during the winter[4] It seems clear, therefore, that 'nuclear winter' is not necessarily a probable outcome of nuclear war, although it is certainly possible The full range of possible impacts can never be completely narrowed because we cannot have access to the war plans of the nations of the world, nor predict the course of any given war once it began I have shown, however, that for the model considered here, that is, one in which 40–80% of the total distributed urban fuels are burned, further research is needed in order to be able to predict the effects on climate

1 Crutzen, P J & Birks, J W *Ambio* **11**, 114–125 (1982)
2 Turco, R. P, Toon, O B, Ackerman, T P, Pollack, J B & Sagan, C *Science* **222**, 1283–1292 (1983)
3 National Research Council *The Effects on the Atmosphere of a Major Nuclear Exchange* (National Academy Press, Washington, 1985)
4 Pittock, A B *et al. Environmental Consequences of Nuclear War* Vol 1 (Wiley, Chichester, 1986)
5 Bing, G *Lawrence Livermore natn Lab Rep* No UCRL-93192 (1985)
6 Ishikawa, E & Swain, D L *Hiroshima and Nagasaki, The Physical, Medical, and Social Effects of the Atomic Bombings* (Basic Books, New York, 1981)
7 Levi, B G & Rothman, T *Physics Today* **38**, 58–65 (1985)
8 Reitter, T A, Kang, S-W & Takata, A N *Lawrence Livermore natn. Lab Rep* No UCRL-53647 (1985)
9 Simonett, D S, Barrett, T N, Gopal, S, Holsmuller, F J & Sun, G preprint (University of California, Santa Barbara, 1986)
10 Crutzen, P J, Galbally, I E & Bruhl, C *Clim. Change* **6**, 323–364 (1984)
11 Bankston, C P, Cassanova, R. A, Powell, E A. & Zinn, B T *US natn. Bur Standards Rep* No NBS GCR-78 147 (1978)
12 Tewarson, A preprint (Factory Mutual Research, Norwood, Massachusetts, 1984)
13 Carrier, G, Fendell, F F & Feldman, P *Combust Sci. Technol.* **39**, 135–162 (1984)
14 Goldsmith, P, May, F G & Wiffen, R. D *Nature* **210**, 475–477 (1966)
15 Penner, J E & Porch, W M *Atmos Envir* (in the press)
16 Penner, J E, Haselman, L C Jr & Edwards, L L *J Clim. appl. Met.* **25**, 1434–1444 (1986)
17 Cotton W R *Am Scient* 73, 275–280 (1985)
18 Penner, J E & Edwards, L L in *Proc. Conf Cloud Physics* **2**, 83–86 (Am meteor, Soc, Massachusetts, 1986)
19 Pruppacher, H R. in *Proc. int Semin. Nuclear War* (eds Newman, W S & Stipcich, S) 163–174 (Servizio Documentazione dei Lab natn di Frascati dell'INFN, Erice, 1985)
20 Golding B W Goldsmith P Machin N A & Slingo A *Nature* 319 301–303 (1986)
21 Hobbs, P V, Radke, L F & Hegg, D A in *Proc. ICSU-SCOPE Workshop Nuclear Winter Scenario Current Studies of Scavenging Models* (International Committee on Cloud Physics, Tallinn, Estonia, in the press)
22 Malone, R. L, Aver, L H, Glatzmaier, G A, Wood, M C & Toon, O B *J geophys Res* 91, 1039–1053 (1986)

Occultation determination of Neptune's oblateness and stratospheric methane mixing ratio

E. Lellouch*, W. B. Hubbard†, B. Sicardy*‡, F. Vilas§ & P. Bouchet‖

* Observatoire de Meudon, 92190 Meudon Principal Cedex, France
† Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA
‡ Université de Paris 7, France
§ NASA Johnson Space Center, Houston, Texas 77058, USA
‖ European Southern Observatory, La Silla, Chile

The occultation of a star by Neptune on 20 August 1985 was observed at 2.2 μm wavelength with telescopes at the European Southern Observatory (ESO) and the Cerro Tololo Inter-American Observatory (CTIO). The detection of a 'central flash' midway between immersion and emersion has allowed the determination of Neptune's oblateness, ε, and the atmospheric extinction at 2.2 μm, which is related to the stratospheric methane mixing ratio. We find $\varepsilon = (2.08^{+0.19}_{-0.18}) \times 10^{-2}$ *and, assuming a stratospheric temperature of 120 K, infer a value of 0.6% (with an uncertainty of a factor of 10) for the methane mixing ratio* (CH_4/H_2) *at 0.3 mbar. The latter value may indicate supersaturation of methane in Neptune's stratosphere.*

STELLAR occultations provide an opportunity to study upper planetary stratospheres and to determine planet oblatenesses, they also allow the detection of ring or arc systems around giant planets. Here we report observations of the occultation by Neptune of star no. 39 in the list by P. D. Nicholson, K. Matthews and G. Gilmore (personal communication) on 20 August 1985, made with the ESO 1-m telescope and the CTIO 1.5-m telescope. Both observations were made in the K band and with a sampling frequency of 100 Hz. An unexpected highlight of this experiment was the observation of a 'central flash' midway between immersion and emersion, when the star was behind the centre of Neptune. This event allows the independent determination of Neptune's oblateness and stratospheric methane mixing ratio. Here we focus only on the study of the flash. Isothermal fits and inversion of the immersion and emersion curves will be presented elsewhere, together with details of an arc structure detected around Neptune at Mauna Kea during the same night.

Observations

Figure 1 shows the occultation light curves and the structure of the observed flashes. A strong peak (~20% of unocculted stellar flux at ESO and ~15% at CTIO) appears first, followed by a wide plateau with several bumps and then a double-peaked structure, after which the signal decreases. The structures observed at the two sites are well correlated (see also Fig. 5).

Theory and modelling of the flash

Theory. The central flash, produced when the star was behind the central part of Neptune, is due to curvature of the limb perpendicular to the light of sight, which causes convergence of the light to the central area of the geometric shadow of the planet[1]. If, as a first approximation, we consider the planet to be spherical, we see that when the star is exactly behind the centre of the disk, the atmospheric refraction is radially symmetrical: the atmosphere is a spherical lens which focuses the light rays on its axis. For on-axis positions, the focusing of the rays is more efficient than spreading due to differential refraction, and this results in a sudden increase of the intensity (Fig. 2).

Because Neptune is significantly oblate, the limb of the shadow is approximately elliptical, and the atmosphere acts as an elliptical lens. The locus of perfect focusing is no longer restricted to the axis of the lens, but its projection in a plane parallel to the plane of the shadow is a curve called the evolute (or the caustic) of the ellipse. The light rays that fall near a given point on the limb converge to the centre of the curvature of the ellipse at the point concerned, and the caustic is then formed by the centres of curvature of the ellipse (see Fig. 15 of ref. 1).

Letting r_e and r_p be the equatorial and polar radii of the planet at the level probed by the flash, and B ($=-25.04°$) the planetocentric declination of the Earth, the equation of the limb of the shadow is

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \tag{1}$$

where a, the semi-major axis, is equal to r_e, and b, the semi-minor axis, is given by

$$b=\frac{r_e r_p}{\sqrt{r_p^2 \sin^2 B + r_e^2 + \cos^2 B}} \tag{2}$$

The apparent oblateness of the shadow is then

$$\varepsilon' = \frac{a-b}{a} = 1-(1-\varepsilon)(\cos^2 B + (1-\varepsilon)^2 \sin^2 B)^{-1/2} \tag{3}$$

where

$$\varepsilon = \frac{r_e - r_p}{r_e} \tag{4}$$

is the true ellipticity of the planet.

Modelling. To calculate the theoretical profile of the flash, we assume the atmosphere to be isothermal in the region probed by the flash (in particular, any horizontal gradient is neglected) and we use the following method[1]. For a point $M(x, y)$ within the shadow, the intensity $\phi(x, y)$ (normalized to the unocculted flux) is the sum of the contributions from the two to four M_i points of the limb whose normals intersect $M(x, y)$. The contribution from each point M_i is evaluated by solving the Baum and Code[2] equation with a scale height $H = RT/mg$ (where R is the gas constant, T is the temperature, m is the mean molar mass and g is the acceleration of gravity) and, to account for the curvature of the limb, the flux thus obtained is multiplied by the ratio of the radius of curvature of the ellipse at M_i to the distance $d = MC_i$, where C_i is the centre of curvature of the ellipse at M_i. As expected for caustics, the intensity becomes infinite on the evolute if one does not take into consideration the finite size of the star (in this case, 6.5 km radius at Neptune's distance, as estimated from the occultation by the Neptune arc observed from the Canada France Hawaii Telescope (B. Sicardy *et al.*, in preparation). To avoid tedious calculations, the profile of the flash is smoothed in one dimension only, along the most sensitive direction—that is, perpendicular to the evolute. Finally, to allow for atmospheric absorption along the path of the rays, we multiply the whole flash profile by a transmission factor $e^{-\tau}$. The intensity throughout the central flash profile is approximately proportional to the scale height near the ray's closest approach to the planet, the effect on the profile of a transmission factor $e^{-\tau} < 1$ is therefore indistinguishable from the effect of a reduction of the scale height by the same factor. In the absence of multiwavelength data, it is necessary to use independent information about the scale height in the region probed by the flash to decide whether an inferred value of the transmission factor represents a finite τ or instead a lower temperature.

Atmospheric level probed by the flash. For an isothermal atmosphere above the probed level, the refractivity of the atmosphere at the level probed by the flash is given by[3]

$$\nu_f = \frac{\theta(h_f)H^{1/2}}{(2\pi r_e)^{1/2}} \qquad (5)$$

where $\theta(h_f) = r_e/D$ is the light-ray bending angle at the altitude probed by the flash (h_f), and D is the Earth-planet distance (4.44×10^9 km). The number density at the altitude probed by the flash is given by

$$n_f = \frac{L\nu_f}{\nu_{STP}} = \frac{L}{\nu_{STP}}\frac{1}{D}\sqrt{\frac{Hr_e}{2\pi}} \qquad (6)$$

where L is the Loschmidt number ($L = 2.687\times10^{19}$ molecules cm^{-3}), and ν_{STP} is the refractivity of the atmosphere at standard temperature and pressure. For $g = 1{,}090$ cm s^{-2}, and an atmosphere of 90% H_2 and 10% He, we obtain $m = 2.2$ g, so $H(\text{km}) = T(\text{K})/2.88$, and $\nu_{STP} = 1.26\times10^{-4}$ at 2.2 μm (ref. 3). The pressure at the probed level is then $p_f = n_f kT$, where k is the Boltzmann constant.

The column density N_f of the atmosphere traversed by the light rays is

$$N_f = n_f\sqrt{2\pi r_e H}\ \text{molecules cm}^{-2} = \frac{Hr_e}{\nu_{STP}D}\ \text{km atm} \qquad (7)$$

with all the distances expressed in kilometres. Here $N_f = 1.564\times10^{-2}\ T$ km atm.

Fitting the flash profile

Oblateness of Neptune. Once the astrometry is fixed, the shape and height of the flash are determined by the four parameters r_e, ε, H and $e^{-\tau}$. Both the radius and the oblateness of the planet affect the profile of the flash by changing the size of the evolute (its typical dimension is $2\varepsilon' r_e$) and thus its position relative to the path of the telescope in the shadow. The flash profile is much more sensitive to ε than to r_e, and the uncertainty in ε (~20%[4,5]) is much greater than that in r_e (0.2%). We therefore fix r_e at 25,200 km in the calculations.

Occultation measurements have provided temperatures of ~150 K at the 1 μbar level[5-7], and far-infrared photometry indicates that the temperature is ~60 K at the 40 mbar level[8,9]. Appleby's[10] radiative convective models show that the temperature can vary from 70 to 150 K at the 0.1-1 mbar level, depending on the amount of aerosol heating and possible CH_4 supersaturation, but also that the temperature at this level must be >100 K to account for the high brightness temperatures observed at 20 and 35 μm. For our 'nominal' model, we adopted $T = 120$ K, the corresponding pressure is then 0.32 mbar.

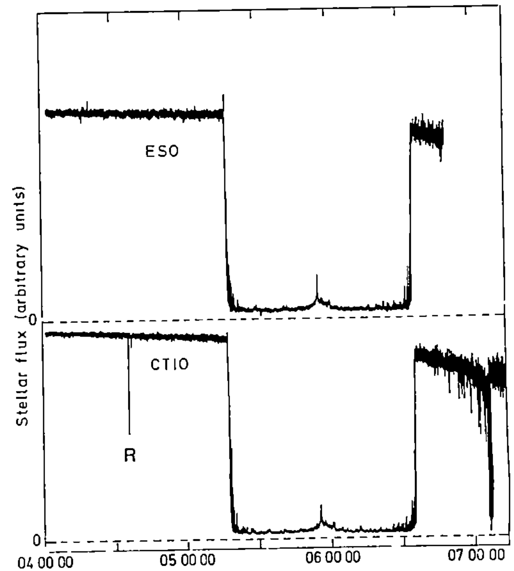


Fig. 1 Occultation light curves and the central flash as observed from ESO and CTIO. Each light curve is plotted at 1-s time resolution. The excursion labelled 'R' is caused by re-centering of the telescope.

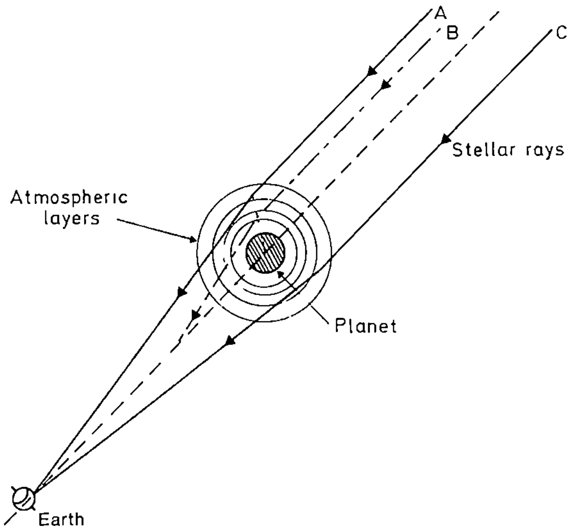


Fig. 2 Schematic explanation of the central flash for a spherical planet. Rays that traverse the atmosphere at different depths, such as A and B, are spread by differential refraction, in contrast, two rays which traverse the atmosphere at the same depth (such as A and C) converge to the same point on the axis and produce the flash.

We started with prediction astrometry provided by P. D. Nicholson (personal communication), which was derived from published occultation mid-times and Neptune shape parameters[4,5]. We used Neptune ephemeris data from the Annuaire du Bureau des Longitudes (1985) (P. Bretagnon, personal communication). To allow for the uncertainty in this initial astrometric solution, two additional parameters were left free: a possible shift Δs between the predicted and actual position of the observatory along the motion of the shadow, and the 'impact parameter' p_0, which is the distance of closest approach to the centre of the shadow.

Our method for fitting the flash was as follows. For a reasonable set of parameters ($\varepsilon = 0.021$, $T = 150\ K$, $e^{-\tau} = 1$, position angle of the projected north pole $P_n = 22.9°$, as determined from the angular momentum of the Neptune-Triton system, assuming a Triton to Neptune mass ratio of 0.00128 (ref. 11)), the theoretical profile was calculated, and cross-correlated with the observed profiles, to obtain the central flash times. We obtained $t =$ 05 55 19 UT for the ESO flash and $t =$ 05 55 21.0 UT for the CTIO flash. Using these times, P_n was determined to be 20.1°. A simultaneous least-squares fit to the ESO and CTIO flash profiles with P_n as an additional parameter yielded $P_n = 20.0°$, in good agreement with the results of the first method. As a

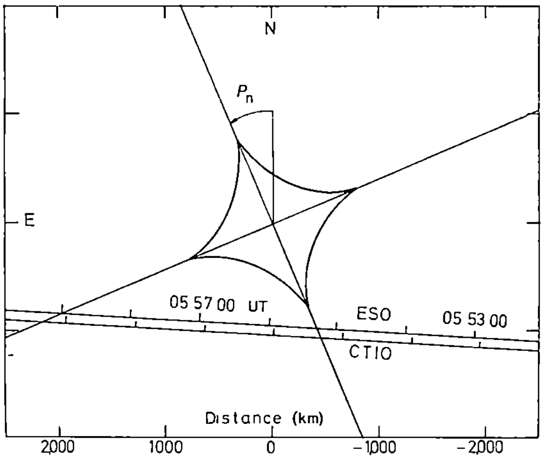


Fig 3 Path of ESO and CTIO relative to Neptune's evolute as projected in the plane of the sky For this diagram, r_e is taken to be 25,200 km, the apparent oblateness of the shadow, $\varepsilon' = 0\ 0171$ and $P_n = 20\ 1°$

further check on the uncertainty in P_n, we measured the times of the peak central intensity at both stations by performing a least-squares fit of a gaussian profile to only those data points near the peak This method yielded $P_n = 19\ 3°$ We estimate the uncertainty in determination of central peak times to be 0 2 s, and the uncertainty in P_n to be 1° Assuming a zero mass for Triton, Harris'[11] ephemeris would yield $P_n = 21\ 6°$ The disagreement with Harris'[11] ephemeris thus appears to be significant, but must be taken with some caution as small atmospheric density fluctuations can cause significant systematic distortions of the central intensity profile The shift Δs was determined to ensure that each observatory crossed the evolute axis at the time of the flash Then, for a grid of points (ε, p_0), theoretical profiles were calculated for a transmission equal to unity, and for each point of the grid, a linear regression was made to find the transmission and the zero level that best matched the observed curve, and the corresponding residuals were calculated The best fit was obtained for a particular set of parameters ε, p_0 and $e^{-\tau}$

Figure 3 shows the paths of the observatories with respect to the evolute in the plane of the shadow, Table 1 summarizes the results for the two observatories As explained above, the factor $e^{-\tau}$ for any assumed temperature (in K) can be scaled from the results presented in Table 1 by multiplying the tabulated values by $120/T$ Note that the agreement between the two observatories is satisfactory Figure 4 shows the quality of the fit in the central part of the flash Our determination of Neptune's oblateness at ~0 3 mbar is $\varepsilon = 0\ 0208^{+0\ 0019}_{-0\ 0018}$ The location of the centre of the planet, as deduced solely from our model fit to the central flash, is in excellent agreement with the oblateness and centre location determined from the limb profile fitted solely to the edge occultation points (discrepancy in centres is $<H$, data to be published elsewhere)

As shown in Fig 5, the model predicts a weak secondary peak at ~05 58 UT, caused by the approach of the occultation tracks to the eastern tip of the evolute pattern (Fig 3) Although the data show a general plateau in this region, secondary peaks in the data do not appear at the predicted time Adjusting the model so as to force a detailed fit only to the secondary peaks violates other constraints Thus, for example, it becomes impossible to fit the main peak correctly, and the location of the planet's centre must be moved away from agreement with the limb profile centre It is possible to argue (to be presented elsewhere), that the secondary peaks in the central flash region are caused by modest deviations of atmospheric refractivity from the smooth, symmetric model assumed here, and that these peaks are produced by essentially the same type of atmospheric disturbances responsible for the well-known rapid intensity fluctuations seen during immersion and emersion Thus a least-squares fit to a model representing the mean properties of the atmosphere should yield valid results, although it will not fit the data in every detail In particular, because scattering by refractivity fluctuations must conserve energy, the mean intensity level in the entire flash region will be largely unaffected by the fluctuations, and the inferred value of $e^{-\tau}$ will be correspondingly conserved

Methane mixing ratio in Neptune's stratosphere. The inferred opacity along the path followed by the light rays is related to Neptune's 2 2-μm absorbers Note that Rayleigh scattering provides negligible extinction (10^{-4}–10^{-5}) at 2 2 μm for paths of a few km atm (ref 12), as is also the case for scattering by fine aerosols[12,13], and that at the pressure probed by the flash (<0 5 mbar), H_2-H_2 and H_2-He pressure-induced absorptions are negligible Thus, the observed absorption must be attributed entirely to methane

In the region of transmission of the filter ($\lambda = 2\ 18$ μm, $\Delta\lambda = 0\ 42$ μm) the CH_4 rotation-vibration bands are those of the so-called octade Three of the eight bands of the octade ($\nu_1 + \nu_4$, $\nu_3 + \nu_4$, $\nu_2 + \nu_3$) have been well studied, the others being much weaker[14-16] Only these three bands were considered here For each of the three bands, each P, Q and R sub-branch shows J-multiplets Integrated intensities of each multiplet up to $J \simeq 14$ were provided by the Laboratoire de Spectronomie Moléculaire at the Université de Dijon (J -C Hilico, personal communication), together with the number of lines involved in each multiplet Because of the large path (~2 km atm) we had to consider even weaker lines (up to $J \simeq 18$), whose intensities were evaluated by assuming they are governed by the exponential term, $\exp(-hcE/kT)$ (where h is Planck's constant, c is the speed of light and E is the lower energy level of the transition) Because of the quickly increasing number of lines in each J-multiplet, we used for each P, Q and R J-multiplet of each band, a band model of non-overlapping independent lines of unequal intensities, assuming an exponential distribution of the intensities[17] For a given multiplet, where σ is the mean line intensity, the probability that a line has an intensity between S and $S + dS$ is $\sigma^{-1} e^{-S/\sigma}\, dS$, and the mean transmission over the band is $T = \exp(-\bar{A})$, where

$$\bar{A} = \frac{1}{\delta} \int_0^{\infty} e^{-S/\sigma}\, dS \int_{-\infty}^{\infty} [1 - \exp(-Sf_\nu a)]\, d\nu \qquad (8)$$

in which δ is the mean spacing of the lines, a is the column density of the absorber and f_ν is the line profile In our case,

Table 1 Results of the isothermal fit

	ESO	CTIO
Time of flash (UT)	05 55 19 1	05 55 21 0
Sky-plane coordinates at time of flash*		
ζ_g (km)	−5,746 7	−5,726 5
η_g (km)	1,378 0	1,379 3
ζ_p (km)	5,395 3	5,344 0
η_p (km)	−2,341 0	−2,424 6
Shift Δs† (km)	—	1
Shift Δp_0‡ (km)	—	−2 5
Oblateness	$0\ 0203^{+0\ 0011}_{-0\ 0013}$	$0\ 0212^{+0\ 0015}_{-0\ 0017}$
Transmission§	$0\ 72^{+0\ 18}_{-0\ 14}$	$0\ 67^{+0\ 21}_{-0\ 16}$
Methane mixing ratio§		
Upper limit	2×10^{-2}	4×10^{-2}
Lower limit	1×10^{-4}	2×10^{-4}
Most probable value	4×10^{-3}	8×10^{-3}

* Subscript g refers to geocentric coordinates, and subscript p refers to displacements due to parallax effects

† With respect to ESO shift

‡ Correction to impact parameter, with respect to ESO correction

§ Assuming a stratospheric temperature $T = 120$ K

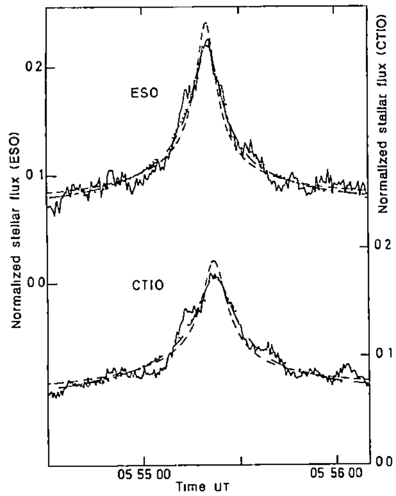


Fig. 4 Fitting of the central part of the flash for various values of ε For ESO ———, ε = 0 0214, – – – –, ε = 0 0203, , ε = 0 0190 For CTIO ———, ε = 0 0227, – – – –, ε = 0 0212, , ε = 0 0195 The data (——) are plotted at a time resolution of 0 5 s

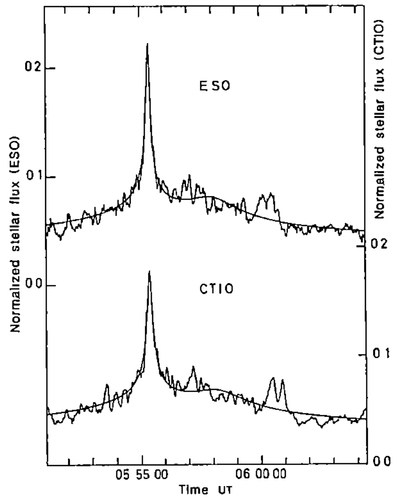


Fig 5 Fitting of the entire flash profile For ESO, ε = 0 0203, for CTIO, ε = 0 0212 The data are plotted at a time resolution of 2 s

the Doppler width $\alpha_D = 4 3\times10^{-7}\ \nu\sqrt{T/M}$ (where M is the molar mass of CH_4) is much greater than the Lorentz width α_L (typically $\alpha_D = 6\times10^{-3}$ cm^{-1} versus $\alpha_L = 3\times10^{-5}$ cm^{-1}), and the cores of the lines have a Voigt profile[18], the far wings being Lorentz-shaped

For each multiplet, the integrated intensity and the number of the lines provides σ and δ Finally, for each temperature and various methane mixing ratios, we obtained the general transmission over the filter as the product of the transmission of each multiplet

The methane mixing ratio was adjusted to match the observed opacity, and our conclusions are summarized in Table 1 If $T = 120$ K (nominal model) we find a methane mixing ratio of 0 6% with an uncertainty of a factor of 10 If $T = 150$ K, the methane mixing ratio is 1% with a factor of 4 uncertainty If $T = 100$ K, we can derive only an upper limit of 2%, with 0 1% as the most probable value

Discussion

Oblateness. Measurements of planetary oblateness are of interest because they are related to internal structures through the gravitational fields To first order, the following relation holds[5]

$$\varepsilon = \tfrac{1}{2}J_2(3 + 1/\Lambda_2) \qquad (9)$$

where J_2 is the gravitational quadrupole moment, and Λ_2 is the second-degree response coefficient as defined by Hubbard[19] the smaller the value of Λ_2, the more centrally condensed is the planet

Recent determinations of ε were obtained from the results of the 15 June 1983 occultation by Neptune From separate analyses of astrometric solutions for various stations, Hubbard *et al*[5] derived $\varepsilon = 0 022 \pm 0 004$, whereas French *et al.*[4] found $0 0191 \pm 0 0017$ We find here an intermediate value of $\varepsilon = 0 0208^{+0 0019}_{-0 0018}$, which is in fairly good agreement with these two determinations It must be emphasized that, due to the very unusual path of the observatories in the shadow (they passed very close to the evolute but did not cross it), the shape of the flash is very sensitive to the oblateness of the model, which accounts for the small uncertainty in our determination of ε

J_2 is equal to 0 0037-0 0043, as determined from the motion of Triton[11] Such values yield $\Lambda_2 = 0$ 12-0 15, which would imply that Neptune is less condensed than Uranus ($\Lambda_2 = 0 09$)

Methane mixing ratio Since the discovery of emission features due to methane at 7 μm in the spectrum of Neptune[20], there have been strong indications that methane could be supersaturated in Neptune's stratosphere Even before this observation, Macy and Trafton[21] had suggested methane supersaturation from brightness temperatures in the 20-μm region, and had found a stratospheric methane mixing ratio of 2×10^{-3} Courtin *et al*[22] also derived a methane mixing ratio of 2×10^{-3}, from analysis of the 8-μm spectrum of Neptune However, Gillett and Rieke[23] suggested that the high brightness temperatures at 8 μm can be explained in terms of a high stratospheric temperature instead of methane supersaturation, and they considered the possibility of a methane mixing ratio as low as 10^{-6}, with a stratospheric temperature of 150 K Macy[24] reanalysed these works, and derived a methane mixing ratio of 10^{-2}-10^{-3} in the stratosphere More recently, Orton *et al*[25] reported the first unambiguous detection of Neptune in the 10 3-11 4-μm region and derived a stratospheric methane mixing ratio of 0 01-0 1, probably approximately equal to the tropospheric value Our observation, which for an assumed stratospheric temperature of 120 K implies $CH_4/H_2 = 0$ 1-1%, also suggests a large enrichment of stratospheric methane, compared to the case where methane is trapped at the tropopause (assuming a tropopause temperature of 53 K (ref 9), the stratospheric mixing ratio would be 5×10^{-5}) These results require that methane is migrating through the tropopause cold trap, and several explanations have been proposed[8,25,26] (1) injection of methane into the stratosphere by strong, localized atmospheric convective motions, (2) transport of small methane particles into the warm stratosphere where they re-sublime, upward eddy diffusion overcoming the sinking of particles due to gravity, (3) an atmosphere so stable and homogeneous and nucleation sites so rare that methane remains supercooled in the inversion region, and (4) the existence of substantial areas where the inversion level is much warmer (leaks in the cold trap) Note that Appleby's[27] Neptune model suggests that the stratospheric methane abundance is near local saturation, a mechanically unstable situation that could, by temperature perturbations, produce an 'inverted cloud'

Conclusion

We have here derived values for Neptune's oblateness and stratospheric methane mixing ratio from an isothermal model of the central flash profiles We emphasize that the oblateness is derived solely from the flash profiles, and does not take into account the emersion and immersion times A determination of

Neptune's oblateness from emersion and immersion times will be presented elsewhere, and comparison of the results will be a good test for the two methods Nevertheless, the fact that the times and shapes of the two flashes observed independently are compatible with the relative positions of the observatories (see Table 1) gives us confidence in the model presented here This model, however, does not fit the complex structure of secondary bumps The effects of non-isothermal wave-like structures in the atmosphere are under study and will be presented elsewhere

E L thanks Jean-Claude Hilico for providing the intensities of the CH_4 bands, W B H and F V acknowledge the support of NASA grant NSG-7045 and thank Brooke Gregory and Patricio Ugarte for essential contributions to the CTIO experiment, and we thank J L Elliot for constructive comments F V is a visiting astronomer at CTIO, supported by NSF contract AST78-27879

Received 3 July, accepted 29 September 1986

1 Elliot, J L *et al. Astrophys J* **217**, 661-679 (1977)
2 Baum, W A & Code, A D *Astr J* **58**, 108-112 (1953)
3 Wasserman, L H & Veverka, J *Icarus* **20**, 322-345 (1973)
4 French, R. G *et al. Astr J* **90**, 2624-2638 (1985)
5 Hubbard, W B *et al. Astr J* **90**, 655-667 (1985)
6 Veverka, J, Wasserman, K. & Sagan, C *Astrophys J* **189**, 569-575 (1974)
7 French, R. G, Elias, J H, Mink, D J & Elliot, J L *Icarus* **55**, 332-336 (1983)
8 Tokunaga, A T, Orton, G S & Caldwell, J *Icarus* **53**, 141-146 (1983)
9 Moseley, H, Conrath, B & Silverberg, R F *Astrophys J* **292**, L83-L86 (1985)
10 Appleby, J F *Icarus* **65**, 383-405 (1986)
11 Harris, A W *NASA Conf Publ.* **2330**, 357-373 (1984)
12 Fink, U & Larson, H P *Astrophys J* **233**, 1021-1040 (1979)
13 Podolak, M & Giver, L P *Icarus* **37**, 361-367 (1979)
14 Brown, L R. *J molec. Spectrosc* **96**, 94-105 (1982)
15 Hunt, R H, Brown, L R, Toth, R A & Brault, J W *J molec. Spectrosc.* **86**, 159-169 (1981)
16 Hilico, J C, Degni, J, Champion, J P & Guelachvili, G *J molec. Spectrosc.* **81**, 277-302 (1980)
17 Goody, R. M *Atmospheric Radiation* (Oxford University Press, 1964)
18 Rao, K N *Molecular Spectroscopy Modern Research* (Academic, New York, 1976)
19 Hubbard, W B, *NASA Conf Publ* **2330**, 291-325 (1984)
20 Macy, W & Sinton, W *Astrophys J* **218**, L79-L81 (1977)
21 Macy, W & Trafton, L M *Icarus* **26**, 428-436 (1975)
22 Courtin, R., Gautier, D & Lacombe, A *Icarus* **37**, 236-248 (1979)
23 Gillett, F C & Rieke, G H *Astrophys J* **218**, L141-L144 (1977)
24 Macy, W *Icarus* **41**, 153-158 (1980)
25 Orton, G S, Tokunaga, A T & Caldwell, J *Icarus* **56**, 147-164 (1983)
26 Orton, G S & Appleby, J F *NASA Conf Publ* **2330**, 89-155 (1984)
27 Appleby, J F thesis, State Univ New York at Stony Brook (1980)

LETTERS TO NATURE

Have burnt-out galaxies and galaxy clusters been detected?

Joseph Silk

Institut d'Astrophysique de Paris, 98 bis, Boulevard Arago,
F-75014-Paris, France and Department of Astronomy,
University of California, Berkeley, California 94720, USA

Close pairs of quasars are receiving much attention for two reasons. The image doubling may be due either to gravitational lensing of a single quasar by an intervening object or to physical clustering, that is, spatially close quasars with similar redshifts. Here I show that an elevated rate of star formation may be responsible for burning out protogalaxies in the densest galaxy clusters, leading both to formation of candidate lenses and to an enhanced probability of close association of quasars in regions of apparently low galaxy density.

Interpretation of identical multiple images of quasars as gravitational lenses requires the lensing objects to be at least as massive as a galaxy[1,2] and in one recent case a galaxy cluster[3], but in several cases to lack the optical luminosity one would ordinarily expect For the most extreme example, the quasar pair[3] 1146+111 B,C with a separation of 157 arc s, the hypothesized lens mass is $\geqslant 10^{15} M_\odot$ In the absence of any very rich galaxy cluster, candidate lenses include a cosmic string[4,5] and a supermassive black hole[6] The recent discovery[7] of spectral differences between the image pair 1146+111 B,C however, suggests that one may in this case be seeing a physically-associated quasar pair This is not easy to understand unless quasars are much more common in large groups or rich galaxy clusters at redshift $z \simeq 1$ than at the present epoch[8] An equally serious problem is the lack of candidate objects of galactic mass for lensing some of the several arc-second separation quasar pairs alternatively, in these cases too, the physical association of quasars must have been enhanced

A possible explanation arises naturally in standard galaxy formation models I will argue that there is a fundamental bias against detecting luminous matter in the densest large-scale regions of the Universe, namely great clusters of galaxies This provides a reasonably natural mechanism both for 'hiding' galaxies to provide candidate lensing objects and for enhancing the frequency of occurrence of the quasar phenomenon in such regions

Consider a situation in which galaxy clustering developed hierarchically after galaxies formed, as in the cold dark matter model[9], although this assumption is not crucial to my argument Vigorous star formation at the present epoch is tidally induced, whether by generation of density waves excited by the tumbling or rotation of a central bar[10] or by the close approach of a companion galaxy[11] Indeed, the most intense bursts of star formation that are identified with IRAS (the Infrared Astronomical Satellite) galaxies are associated with interacting galaxies[12] From the theoretical point of view, the non-linear response of a cold system greatly amplifies even weak tidal perturbations[13] The net effect is to enhance gas cloud collisions and dissipation, initiate large-scale gravitational instabilities, and trigger star formation Empirically, the IRAS studies suggest that the specific star formation rate is enhanced by one or even two orders of magnitude in the most intense star bursts, such as that in Arp 220, relative to the solar neighbourhood

Little is known about star formation in protogalaxies But it is virtually certain that when the old spheroidal stellar component formed, the specific star formation rate was anywhere between two and three orders of magnitude higher than it is today in the solar neighbourhood It was necessary to have formed $\sim 10^{11} M_\odot$ of low-mass stars in a timescale that is estimated to be $\sim 10^9$ yr It cannot be much less, because stellar evolution calculations and chemical evolution models require this time to build up the heavy elements that are seen in the old stars and to account for the spectral synthesis of elliptical galaxies[14] It cannot be much longer, because the oldest stars in our Galaxy show no evidence of any spread in formation epoch[15] It is tempting to identify the initiator of protogalactic star formation activity with the similar trigger that is observed today in starburst galaxies Indeed, strong tidal interactions are inevitable if galaxies formed hierarchically from a Zeldovich spectrum of primordial adiabatic, gaussian density fluctuations, in which small-scale power has not been suppressed, as in the cold dark matter model[9] Moreover, all available evidence suggests that the protogalactic initial stellar mass function (IMF) spanned a similar mass range to the present day IMF[16,17], thereby helping to justify our confidence in taking over inferences about star formation in nearby galaxies

One additional ingredient of star formation has received considerable attention The efficiency of star formation measures the mass fraction of gas that has been converted into stars In the solar neighbourhood, the long-lived, low mass ($\leqslant 1 M_\odot$) stars account for most ($\sim$85%) of the total star formation rate[18] In regions of vigorous star formation, however, including the inner galaxy[19], the situation appears to be rather different. In particular, in starbursts, the ratio of the observed massive star formation rate to the total gas supply implies that relatively few low mass stars can be forming[20]. The IMF cannot extend with any appreciable power below 2-3$M_\odot$, and in some cases, 5-10$M_\odot$. There are plausible, although hardly rigorous, theoretical arguments concerning why this should be so[21,22].

The following inference now becomes logical, and possibly inevitable. If there are any regions of the Universe so dense that tidal encounters between protogalaxies occurred at a rate large enough to stimulate star formation to a level up to and even beyond that seen in starbursts, then few very-low-mass stars could have formed before massive star formation had effectively depleted the net gas supply. If the star-forming phase in a proto-elliptical continues for $\sim 10^9$ yr, there is enough time to form many low-mass stars. But if it lasts only $\sim 10^8$ yr, low-mass stars may barely form at all. The evidence for this is largely circumstantial. Star formation bursts typically last $\sim 10^8$ yr, and fail to form low mass stars, whereas low mass stars do form over a much longer timescale, some several billions of years in the solar neighbourhood[23]. This bimodality between short-lived massive and long-lived low-mass star formation provides an attractive explanation of the local paucity of metal poor G-dwarfs in the disk[24]. Hence I conclude that in any region where the characteristic tidal interaction timescale between galaxies is as short as $\sim 10^8$ yr, one would be in grave danger of exhausting the gas supply without producing a significant number of low-mass stars. In 10^8 yr, most of the mass lost by stellar evolution of massive stars is recycled: if the low-mass cutoff in the IMF is above $\sim 10M_\odot$, essentially all the initial protogalaxy mass is cycled into compact remnants. With an IMF $dN(m)/dm \propto m^{-3}$ for massive stars[25], at most 20% of the initial mass is returned even if the low-mass cutoff is $\sim 5M_\odot$. The bulk of the initial mass is trapped in compact remnants. Evidently, with three or more generations of star formation, one would consume $\gtrsim 90\%$ of the initial gas supply even if the star formation efficiency is $\sim 50\%$ (as is required to form bound star clusters). Any stars of lower mass would presumably not have had time to form in significant numbers because their formation rate is too low. Exactly why bimodal star formation occurs is not fully understood, but there is ample circumstantial evidence, both from our Galaxy and from other galaxies, that massive star formation occurs at a high specific rate without accompanying low-mass star formation, while low-mass star formation proceeds at a more leisurely pace, and is generally accompanied by some massive star formation[19,24]. Alternatively, verification of a flat IMF as may be present in metal-rich galactic globular clusters[26] would show that some bound stellar systems are capable of eventually becoming seriously deficient in visible stars relative to those with a solar neighbourhood IMF[27]. This directly suggests that metal-rich galaxies could burn out after successive episodes of massive star formation have exhausted the gas supply.

Consider the double quasar 1146+111 B,C with a separation of 157 arc s. I first assume that the multiple images are due to a non-singular gravitational lens. Placing the lens at the optimal redshift to minimize its inferred mass, together with the observed separation between images, specifies the gravitational potential well of the lensing object, apart from a geometrical uncertainty[28]. Assuming the lens to be spherical, one infers a surface density $\Sigma \gtrsim 1$ g cm^{-2} and virial velocity dispersion $V_c \simeq 3{,}000$ km s^{-1}. If it is indeed a cluster of galaxies, the galaxy crossing timescale is $V_c/G\Sigma \simeq 10^8$ yr. This dynamical timescale is sufficiently short (a typical value for Abell clusters is $\sim 10^9$ yr) that one may expect tidal interactions to have had an important role in cluster evolution.

To evaluate the role of tidal heating, I use an effective cross-section for energy transfer derived from numerical simulations[29], $\sigma_E = 30r_e^2(V_*/V_c)^3$, where r_e is the effective radius of a galaxy, taken to have a de Vaucouleurs profile and to all be identical, and V_* is the isotropic r.m.s. stellar velocity dispersion within a galaxy. The heating rate is

$$E/E = 1.4(G\Sigma)^2\sigma_E/r_eV_*^2V_c = 0.7\Sigma^2r_3V_{*,300}V_{c,3000}^{-4}(10^9\text{ yr})^{-1}$$

where $r_3 = r_e/3$ kpc, $V_{*,300} = V_*/300$ km s^{-1} and $V_{c,3000} = V_c/3{,}000$ km s^{-1}. The actual situation would probably have been even more extreme, because during the initial collapse of the cluster the galaxy velocity was $\ll V_c$ and protogalaxies had considerably larger effective radii. Evaluation of E/E $(= n_c\sigma_EV_c)$ for parameters typical of the Coma clusters of galaxies ($n_c \simeq 200$ Mpc^{-3} in the core[30]) yields, for example, only $E/E \simeq 0.01(10^{10}\text{ yr})^{-1}$. Thus E/E is some three orders of magnitude larger for the hypothesized cluster that is lensing 1146+111 B,C.

In the case of cold dark matter with primordial density fluctuations described by the usual power spectrum $|\delta_k|^2 \propto k^n$, I find that for scales just going nonlinear at redshift z, $E/E \propto (1+z)^{2(n+5)/(n+3)}$, or over rich cluster scales where $n \simeq -1$ for a Zeldovich fluctuation spectrum, it follows that $E/E \propto (1+z)^4$ provided $V_* < V_c$. The density of the Coma cluster core corresponds to turn-around at background density equal to about $(1/200)\rho_c \simeq 10^{-27}$ g cm^{-3}, where ρ_c is the observed core density, or at $1+z \approx 6\Omega^{-1/3}$ if Hubble's constant $H_0 = 50$ km s^{-1} Mpc^{-1}. Hence galaxy formation at $1+z \gtrsim 10$ would have resulted in producing burnt-out galaxies only when clusters rather rarer (earlier forming) than Coma first collapsed. Abell clusters contain $\sim 5\%$ of luminous galaxies, and their frequency corresponds to their being $\sim 2\sigma$ fluctuations in a primordial gaussian fluctuation spectrum appropriate to cold dark matter models[31]. Clusters which would have induced burn-out are accordingly $\sim 3\sigma$ fluctuations, since $(\delta\rho)_c \simeq (1+z_c)(\delta\rho/\rho)_{r.m.s.}$, and should be relatively rare, amounting at most to ~ 1 per every 5 Abell clusters over the same filtering scale (the number density of density peaks being proportional to[32,33] $\nu^2\exp(-\nu^2/2)$ for $\nu\sigma$ fluctuations if $\nu \gg 1$).

I conclude that galaxies in clusters as extreme as that inferred to be producing 1146+111B,C would have undergone extremely strong tidal interactions during the initial collapse of the cluster, as well as subsequently. Such clusters should be rare compared with Abell clusters, but still sufficiently frequent to lens quasars with reasonable probability. Galaxy burn-out may even be a common enough phenomenon to account for some of the lenses which produce quasars pairs with arc-second separations: a definite prediction of my model is that there should also be a rich cluster of burnt-out galaxies in such cases, so that broad separation ($\gtrsim 100$ arc s) doubling should be sought in the fields of putative lenses where no luminous galaxy candidate has hitherto been found.

Next, I consider the possibility that formation of burnt-out galaxies may also be associated with an enhanced frequency of occurrence of quasar pairs. The tidal interactions that are able to stimulate massive star formation are also capable of driving molecular clouds and stars onto radial orbits which fuel the massive black hole that may already have formed at the centre of a galaxy. Precisely when or how such massive holes form is a contentious issue, but it is generally believed that the fuelling of such a hole of $\sim 10^8 M_\odot$ at a rate $\gtrsim 10\ M_\odot$ yr^{-1} is responsible for the quasar phenomena. Indirect evidence from absorption line studies suggest that quasars at $z \gtrsim 1.2$ are indeed in rich clusters[34,38], while direct imaging reveals that low-redshift quasars apparently avoid rich clusters[39]. As quasars evolve strongly in luminosity with redshift, this provides circumstantial evidence that the rich cluster environment may enhance quasar activity. The burnt-out galaxy hypothesis may render many of the associated galaxies sub-luminous, others, however, may be caught in their luminous protogalactic starburst phase. It is intriguing that the probable narrow emission line galaxy[35] at $z = 3.218$ that is a companion to the quasar PKS1614+051 is unusually luminous for a normal galaxy, and is perhaps the best candidate to date for a protogalaxy with the characteristic strong

Ly-α emission and compact appearance that models[36] predict

Finally, I note that one prediction, if a rich galaxy cluster is responsible, either by locally enhancing the spatial frequency of quasars or by lensing, for the quasar pair 1146+111 B,C, is that gas infall and ensuing heating could lead, at least in the case of the lens hypothesis, to a potentially detectable Sunyaev-Zeldovich microwave background anisotropy[37] and, more seriously, to excessive hard X-ray emission, J P Ostriker and J S have estimated (unpublished) that the thermal bremsstrahlung X-ray flux should be $\sim 10^{11}$ $(f_b/0\ 1)^2$ erg cm^{-2} s^{-1} keV^{-1} up to $\sim kT \simeq 30$ keV, where f_b is the mass fraction of the cluster in diffuse gas, and is probably well above the HEAO-1 all sky survey threshold if $f_b \simeq 0\ 1$

Burnt-out galaxies avoid this difficulty, because the available gas supply was mostly exhausted in forming massive stars Any left-over gas would have participated in a cluster cooling flow, the high density of which should have led to strong cooling in the cluster core and galaxy formation in the same tidally stirred-up environment where the burnt-out galaxies previously formed There should be some residual gas, however, at a density low enough that cooling should be unimportant (that is $f_b \simeq 0\ 01$) This means that up to $\sim 10^{13} M_\odot$ of hot gas could be present, and it would contain processed matter, including dust Thus the search for diffuse matter in the dark cluster that may be responsible for the quasar pair 1146+111 B,C should be pursued at X-ray, microwave and infrared wavelengths More generally, the rarity of the burnt-out clusters guarantees that any contribution to the cosmic background light should be negligible However, the possibility of observing vigorous starburst activity in the vicinity of high-redshift quasars deserves further attention This activity could manifest itself, for example, in producing extended emitting or absorbing clouds of low ionization state and with velocity dispersions of up to $\sim 1{,}000$ km s^{-1} that are characteristic of vigorous starburst activity

This research has been supported in part by NASA and by NSF I thank Jean Audouze for his hospitality at I A P where this work was completed

Received 2 June, accepted 28 August 1986

1 Djorgovski, S & Spinrad, H *Astrophys J* **282**, L1-4 (1982)
2 Weedman, D, Weymann, R. J, Green, R. F & Heckman, T M *Astrophys J* **255**, L5-9 (1982)
3 Turner, E *et al. Nature* **321**, 142-144 (1986)
4 Paczynski, B *Nature* **319**, 567-568 (1986)
5 Gott, J R *Nature* **321**, 420-421 (1986)
6 Paczynski, B *Nature* **321**, 419-420 (1986)
7 Shaver P & Christiani, C *Nature* **321**, 585-586 (1986)
8 Bahcall, J N, Bahcall, N A & Schneider, D P preprint (Institute for Advanced Study, 1986)
9 Blumenthal, G, Faber, S, Primack, J & Rees, M J *Nature* **311**, 517-525 (1984)
10 Combes, F & Gerin, M *Astr Astrophys* **150**, 327-338 (1985)
11 Icke, V *Astr Astrophys* **144**, 115-123 (1985)
12 Joseph, R D & Wright, G S *Mon Not R. astr Soc.* **214**, 87-95 (1985)
13 Goldreich, P & Lynden-Bell, D *Mon Not R. astr Soc.* **130**, 125-158 (1965)
14 Bruzual, G *Astrophys J* **273**, 105-127 (1983)
15 Sandage, A *Astrophys J* **252**, 553-573 (1982)
16 Silk, J *Mon Not R. astr Soc.* **205**, 705-718 (1983)
17 Cayrel, R *Astr Astrophys* (in the press)
18 Miller, G & Scalo, J *Astrophys J Suppl.* **41**, 513-547 (1979)
19 Gusten, R & Mezger, P *Vistas Astr* **26**, 159-224 (1983)
20 Rieke, G H, Lebofsky, M J, Thompson, R I, Low, F J & Tokunaga, A K. *Astrophys J* **238**, 24-40 (1980)
21 Larson, R *Mon Not R. astr Soc.* **214**, 379-398 (1985)
22 Silk, J *Luminous Stars and Associations In Galaxies, IAU Symp* **116**, 301-315 (Reidel, Dordrecht 1986)
23 Twarog, B *Astrophys J* **242**, 242-259 (1980)
24 Larson, R *Mon Not R. astr Soc.* **218**, 409-428 (1986)
25 Scalo, J *Fund Cosm Phys* **11**, 1-278 (1986)
26 McClure, R D *et al. Astrophys J* **307**, L49-53 (1986)
27 Freeman, K D *The Evolution of Galaxies and Stellar Populations* (eds Tinsley, B M & Larson, R B) 133-148 (Yale University Observatory, 1977)
28 Turner, E, Ostriker, J P & Gott, J R *Astrophys J* **284**, 1-22 (1984)
29 Aguilar, L & White, S D M *Astrophys J* **295**, 374-387 (1985)
30 Bahcall, N *Astrophys J* **247**, 787-791 (1981)
31 Kaiser, N *Astrophys J* **273**, L17-L20 (1983)
32 Doroshkevich, A *Astrophysics* **6**, 320-330 (1970)
33 Bardeen, J, Bond, J R., Kaiser, N & Szalay, A *Astrophys J* **304**, 15-62 (1986)
34 Weymann, R J, Carswell, R. F & Smith, M G *A. Rev Astr Astrophys* **19**, 41-76 (1981)
35 Djorgovski, S, Spinrad, H, McCarthy, P & Strauss, M A *Astrophys J* **299**, L1-5 (1985)
36 Meier, P L *Astrophys J* **207**, 343-350 (1976)
37 Ostriker, J P & Vishniac, E *Nature* **322**, 804 (1986)
38 Foltz, C B *et al. Astrophys J* **307**, 504-534 (1986)
39 Yee, H C K & Green, R. F *Astrophys J* **280**, 79-90 (1984)

A radio nebula associated with Circinus X-1

R. F. Haynes[*], M. M. Komesaroff[*], A. G. Little[†][‡], D. L. Jauncey[*], J. L. Caswell[*], D. K. Milne[*], M. J. Kesteven[*], K. J. Wellington[*] & R. A. Preston[§]

[*] Division of Radiophysics, CSIRO, PO Box 76, Epping, New South Wales 2121, Australia
[†] School of Physics Department, University of Sydney, Sydney, New South Wales 2006, Australia
[§] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109 USA
[‡] Deceased

The investigation reported here of the variable source Circinus X-1 (Cir X-1) reveals that it is embedded in a nebula of steady radio emission extending over several parsecs, orders of magnitude larger than the binary stellar system responsible for the fluctuating component of emission (detected at X-ray, radio and optical wavelengths). This is in marked contrast to most X-ray binaries, where an envelope of radio emission is conspicuously absent. There are difficulties in explaining the emission, but analogies with SS433 and the Crab nebula suggest possible models.

X-ray[1], optical[2], infrared[3,4] and radio studies[2] indicate that Cir X-1 is an eccentric double-star system, with one component of the binary being very compact, possibly a neutron star or a black hole The 16 5-day periodicity, manifest across the observed frequency spectrum[5], implies an orbital semimajor axis $\leqslant$0 5 AU H I absorption measurements[6] yield a firm lower limit to the distance of 8 kpc At this distance the binary orbit subtends an angle $\leqslant 10^{-4}$ arc s, below the resolution limit of even very long baseline interferometry (VLBI) Nevertheless, Preston *et al*[7], using VBLI, reported a quiescent component of emission with angular size $\geqslant$0 2 arc s at 2 3 GHz and Haynes *et al*[2] reported evidence of an extended radio source at 1 4 GHz

The measurements presented here were made with the Molonglo Observatory Synthesis Telescope (MOST), which operates at 843 MHz, at the declination of Cir X-1 a 12-h synthesis observation[8] yields a beam of 43×52 arc s A contour diagram of Cir X-1 and the surrounding region is shown in Fig 1

The fact firmly established by the present results is that the variable compact source Cir X-1 is embedded in a radio nebula extending over ~5 arc min east-west and ~10 arc min north-south which shows no variability The integrated flux of the nebula at 843 MHz is 0 56 Jy

The compact object in Fig 1 at RA 15 h 16 min 48 6 s and dec −56°59′11 4″ (1950 coordinates are used throughout this paper) coincides with the X-ray source[1] Table 1 compares our best estimate of its position with the position measured by Argue and Sullivan[9] for the red star identified with the X-ray source and the position of the compact 2 3-GHz source measured using VLBI[10]

The compact source is a well-known radio variable[2] Between May 1983 and June 1985 at 843 MHz we found that the flux

Table 1 Compact object associated with Circinus X-1

Object	Position (B1950) RA	Position (B1950) Dec	Ref
MOST compact object	15 h 16 min 48 58±0 7 s	−56°59′ 11 4±0 5″	This paper
Red star	15 h 16 min 48 436±0 022 s	−56°59′ 11 4±0 28″	9
2 3 GHz VLBI object	15 h 16 min 48 37±0 04 s	−56°59′ 11 6±0 3″	10

Uncertainties in the MOST positions are dominated by an assumed 0 5 arc s uncertainty in ionospheric refraction

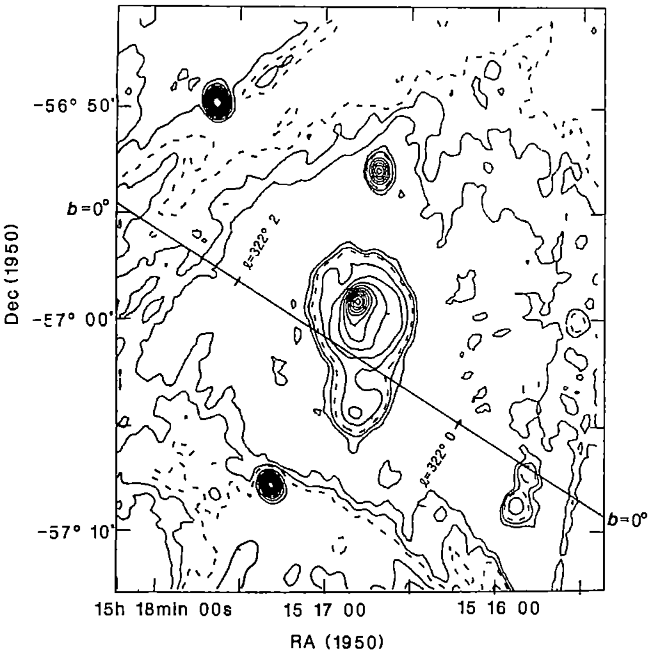


Fig 1 843-MHz map of the Circinus X-1 region obtained by integrating seven 12-h synthesis maps made with the MOST between May 1983 and June 1985 Contours are shown at levels of 1, 3, 5 (dashed), 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 and 90 mJy per beam area

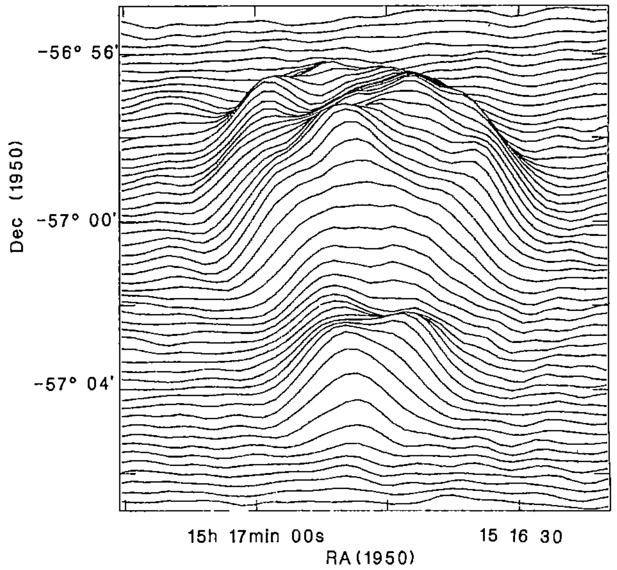


Fig 2 Perspective view of the nebular region surrounding Cir X-1 obtained by subtracting the 843-Hz emission from the point radio source associated with Cir X-1 from the data shown in Fig 1

density varied between 40 and 88 mJy The belief that the X-ray, radio and optical emission all emanate from Cir X-1 ultimately rests on the time variable correlation[2], for example, detailed monitoring in April 1982 showed radio intensity changes clearly correlated with the 3-6-keV X-ray intensity over the same time (L Kaluzienski, personal communication)

To study the detail of the nebula, we have subtracted the point radio source centred on Cir X-1 as shown in Fig 2 The most prominent part of the nebula is approximately centred on the compact component and adjoins a weaker and narrower component to the south centred near RA 15 h 16 min 47 s, dec −57°04′ Earlier, lower-resolution measurements[2,5] showed a quiescent source which we now assume arose predominantly from the diffuse nebula rather than from the compact variable source These measurements may be combined with the present 843-MHz results ($S_{843} = 0\ 56$ Jy) to yield a distinctly non-thermal spectral index of $\alpha \simeq -0\ 5$ ($S \propto \nu^{\alpha}$) Non-thermal sources of this intensity have a spatial density of ~1 per 3 square degree, and Cir X-1 coincides with the nebula's centre to within 1 arc s, the probability of a chance coincidence is then $\leqslant 1$ in 3,000 Thus, we conclude that the whole nebula is physically associated with Cir X-1 rather than a mere line-of-sight coincidence

Cir X-1 has several characteristics in common with Cyg X-1 (refs 11, 12) However, as far as we know, Cyg X-1 displays no extended radio component comparable with that associated with Cir X-1

In attempting to explain the extended distribution we consider two general types of model (1) the nebula is the remnant of an earlier active phase which no longer persists (2) The nebula has its energy supply maintained by continuous injection from a central engine The first interpretation corresponds to the typical supernova explosion In this case, we should expect to find a limb-brightened source There is no evidence of this in Fig 1, even when allowance is made for the limited angular resolution Furthermore, the nebula has a surface brightness that would be anomalously low for a typical shell supernova remnant of this size[13]

In the second kind of model, (2) the Crab nebula represents one variety of continuous injection model In this example, the central engine is a rotating neutron star, and energy transfer to the surrounding nebula may be accomplished by low-frequency electromagnetic waves[14] or by charged particles accelerated to ultra-relativistic energies in the enormous electric fields produced by the induction[15] Either process requires that the central engine be in a highly evacuated cavity Such processes are thus expected to be inoperative in Cir X-1 because it is an eccentric binary system[16,17] with a high rate of accretion from a 'normal' star to a 'compact' star and the resulting mass flow may smother the energy transfer process On the other hand, SS433 (refs 18, 19) another variable binary-star system, produces collimated jets which do successfully escape from the binary system and are not smothered by the mass transfer

The Cir X-1 flares have the characteristics of the van der Laan expanding radio source model[16] This leads us to suggest, as another variant of the continuous injection model, that the extended nebular is the accumulation of the energetic particles and plasma ejected during the flaring activity

Clark *et al* have argued that Cir X-1 is a runaway binary system According to this theory Cir X-1 was ejected from the supernova G321 9−0 3, the centre of which is some 25 arc min south, at the time of its eruption If this could be substantiated, it would rule out the possibility that the nebular surrounding Cir X-1 is itself a supernova remnant, and render a continuous injection model amost inescapable The southward extension of the radio nebular would also fit naturally into such a picture This model requires an asymmetric supernova explosion with a high ejection velocity Accepting 8 kpc as a minimum distance to Cir X-1 and 10^5 yr as the age[12] of G321 9−0 3, we see that the transverse component of the ejection velocity needs to be 500 km s^{-1} This is greater than the largest known pulsar velocity[20] (365 km s^{-1})

There are several future lines of investigation which could resolve some of the questions raised by Cir X-1 First, proper-motion studies in the radio domain (using VLBI techniques) and perhaps in the optical region could determine whether Cir X-1 was ejected from G321 9−0 3 If so the nebula around Cir X-1 could not itself be a supernova remnant We propose to undertake these studies using the Parkes-Tidbinbilla interferometer system Second, further mapping of the nebula could be undertaken to investigate more fully the absence of limb-brightening and the spectral index Finally, detailed theoretical modelling of the dynamics could yield an upper limit on the time for circularization of the eccentric orbit and thus yield an upper limit on the lifetime of the system

The MOST is operated by the University of Sydney under funding provided in part by the Australian Research Grants Committee, the University of Sydney Research Committee and the Science Foundation within the University of Sydney We thank Professor B Y Mills and Dr J M Durdin for help with using the MOST The research described in this paper was carried out in part by the Jet Propulsion Laboratory of the California Institute of Technology under contract to NASA

Received 21 July, accepted 10 September 1986

1 Kaluzienski, L J, Holt, S S, Boldt, E A & Serlemitsos, P J *Astrophys J* **208**, L71-L75 (1976)
2 Haynes, R F *et al Mon Not R astr Soc* **185**, 661-671 (1978)
3 Glass, I S *IAU Circ* No 2974 (1976)
4 Glass, I S *Mon Not R astr Soc* **183**, 335-340 (1978)
5 Whelan, J A J *et al Mon Not R astr Soc* **181**, 259-271 (1977)
6 Goss, W M & Mebold, U *Mon Not R astr Soc* **181**, 255-258 (1977)
7 Preston, R A *et al Astrophys J* **268**, L23-L26 (1983)
8 Mills, B Y *Proc astr Soc Aust* **4**, 156-159 (1981)
9 Argue, A N & Sullivan, C *Observatory* **102**, 4-6 (1982)
10 Argue, A N, Jauncey, D L, Morabito, D D & Preston, R A *Mon Not R astr Soc* **209**, 11P-14P (1984)
11 Braes, L E & Miley, G K *Nature* **232**, 246 (1971)
12 Clark, D H, Parkinson, J H & Caswell, J L *Nature* **254**, 674-676 (1975)
13 Milne, D K *Aust J Phys* **32**, 83-92 (1979)
14 Ostriker, J P & Gunn, J E *Astrophys J* **157**, 1395-1417 (1969)
15 Goldreich, P & Julian, W H *Astrophys J* **157**, 869-880 (1969)
16 Haynes, R F, Lerche, I & Murdin, P G *Astr Astrophys* **87**, 299-302 (1980)
17 Murdin, P G *et al Astr Astrophys* **87**, 292-298 (1980)
18 Clark, D H & Murdin, P G *Nature* **276**, 45 (1978)
19 Margon, B *A Rev Astr Astrophys* **22**, 507-536 (1984)
20 Anderson, B & Lyne, A G *Nature* **303**, 597-599 (1983)

The historical record for Sirius: evidence for a white-dwarf thermonuclear runaway?

Frederick C. Bruhweiler*, Yoji Kondo† & Edward M. Sion‡

* Department of Physics, Catholic University of America, Washington, D C 20046, USA
† Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA
‡ Department of Astronomy and Astrophysics, Villanova University, Villanova, Pennsylvania 19085, USA

Schlosser and Bergmann[1] recently presented evidence that in medieval times Sirius was a bright red star, rather than the present bluish-white star, from which they have suggested that Sirius B is a recently born white dwarf. However, their model poses severe evolutionary problems. We present the results of our attempts to detect possible planetary nebula ejecta toward Sirius, using data obtained by the International Ultraviolet Explorer (IUE) satellite. Based upon these results and in the light of recent advances in understanding white-dwarf evolution, we propose that Sirius B underwent a recent thermonuclear runaway event, triggered by a diffusion-induced CN reaction, to explain the historical behaviour of this star.

Two evolutionary problems remain in the Schlosser and Bergmann evolutionary picture First, the implied cooling time for the white dwarf, Sirius B, is several orders of magnitude too short to be explained by current theories of white-dwarf cooling For example, current predictions indicate that the time required for a white dwarf with an effective temperature $\sim 2\ 7\times10^4$ K, like Sirius B (ref 2), to cool from 10^5 K, typical of a planetary nebula nucleus, is $\sim 10^7$ years To explain such discrepancies would require either new physics or a rethinking of our current picture of the later stages of stellar evolution Secondly, there is no evidence of any stellar ejecta, although such ejecta could have escaped detection A search for such ejecta is warranted, especially in view of recent analysis of ultraviolet spectral data obtained of the red-dwarf-white-dwarf binary system V471 Tauri[3], which showed possible gaseous material moving away from that system at $-590\ \mathrm{km\ s^{-1}}$

It is well known that interstellar ultraviolet absorption lines of cosmically abundant atomic species can provide a very sensitive means of detecting intervening gas Thus, in our search for the putative ejecta, we have made use of data previously acquired by the IUE These data come from the high-dispersion spectral image SWP 10075 obtained for the white dwarf Sirius B These data cover the spectral region from ~1,200 to 2,000 Å and provide information on many ground-state transitions for several cosmically abundant species A full discussion of the acquired data and a discussion of the observed interstellar features can be found elsewhere[4]

The new observational results that we present here represent a reanalysis of the IUE image SWP 10075 This reanalysis is necessary because originally we were not looking for possible high-velocity interstellar absorption, and we can now apply additional techniques to place stronger detection limits on intervening gaseous absorption by using the almost featureless nature of white-dwarf spectra

Our data analysis follows very closely that recently presented[3] We have taken wavelength regions containing the strongest transitions for the abundant atomic species and converted these data to the velocity frame by using the transformation $v = c\,(\lambda - \lambda_0)/\lambda_0$, because the internal wavelength (and velocity) scale of the IUE cameras is extremely accurate ($\pm 2\ \mathrm{km\ s^{-1}}$), absorption due to more than one ion at a particular velocity should be very closely reproduced at any velocity The interstellar features of C II, Si II and O I are seen near zero velocity, but there is no evidence of other features within $3{,}000\ \mathrm{km\ s^{-1}}$ of the rest velocity There was a hint of a feature at the 1 5 σ level near $-30\ \mathrm{km\ s^{-1}}$ with respect to the interstellar component in the Si II data, but nothing is seen in other ions Although planetary nebulae typically show expansion velocities $\sim 30\ \mathrm{km\ s^{-1}}$, the recent discovery of sharp high-velocity features in V471 Tauri justifies searching for components over a large velocity interval Because the species C II, Si II and O I have similar ionizations, we coadded these data in an attempt to enhance the detection limit in case weak high-velocity components of these ions existed We found no evidence of additional absorption other than the expected interstellar features The results of this coaddition are shown in Fig 1 The same test was also performed upon the stronger members of the resonance doublets for the highly ionized species N V, C IV and Si IV Again, no evidence of absorption was found either corresponding to the already observed low-ionization interstellar features or at higher velocities

Based on these results, we find it hard to support the supposition that the Sirius stellar system has recently ejected any type of detectable nebulosity or gaseous shell One might still argue that the velocity resolution ($22\ \mathrm{km\ s^{-1}}$) would not be sufficient if the expanding shell were 'masked' or unresolved from the intervening interstellar features However, the implied average particle densities toward Sirius are in accord with those found in other short lines of sight in the local interstellar medium[4,5] We derive from the IUE data an upper limit of $3\times10^{17}\ \mathrm{cm^{-2}}$ for the total gas column density for the putative ejecta about Sirius B If we assume an age of 1,000 years and an expansion velocity of $30\ \mathrm{km\ s^{-1}}$, we get a corresponding upper limit of $3\ 6\times10^{-5}\ M_\odot$ for the mass of any ejecta In addition, Copernicus results[6], with about a twofold better resolution and much better signal-to-noise ratio, do not indicate multiple interstellar Mg II features at 2,795 and 2,802 Å Thus, we must conclude that the physical mechanism causing the colour change of Sirius was due to a relatively quiescent process that did not involve any significant mass ejection

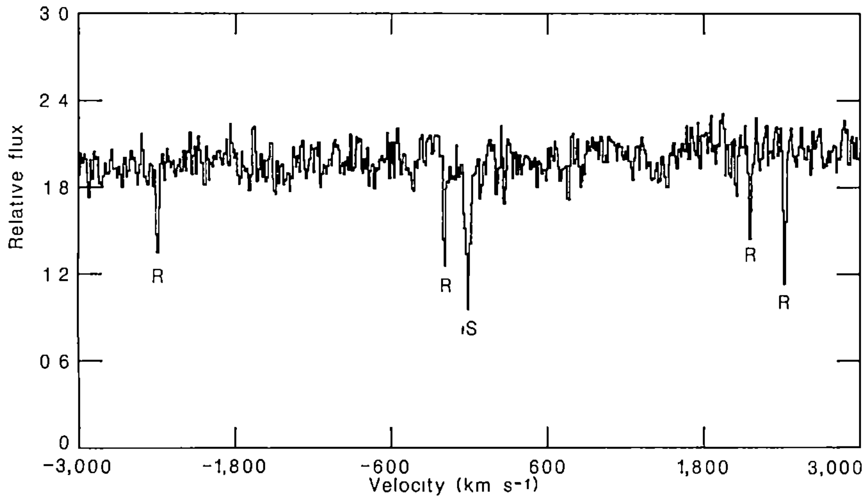


Fig. 1 Coadded data for regions corresponding to strong ground-state transitions of O I, C II and Si II. The coadded data for the O I 1,302.169, C II 1,334.532 and Si II 1,260.421 Å transitions are shown in the velocity frame. No evidence of absorption other than the expected interstellar contribution near zero velocity (labelled IS) is found. The other sharp features (R) are detector fiducial or réseau marks contributing from each of the data sets covering the respective atomic transitions. The data are from IUE image SWP 10075 acquired using the small aperture.

One other possible physical interpretation is that the red colour of Sirius during mediaeval times was due to intervening dust along the line of sight. Even though Sirius does not seem to have an infrared excess, at least two nearby A stars, Vega and Beta Pictoris, show strong evidence of possessing extended dust envelopes or disks[7–10]. The sizes of the dust particles in these disks are inferred to be large, possibly as large as 1 mm. These particles can provide extinction, but do not seem to yield any reddening at optical and ultraviolet wavelengths. Regardless of the properties of the dust, there cannot be reddening without extinction. To obtain an observed $B-V \geqslant 1.0$ as is deduced for Sirius in the sixth century would require substantial dimming. If we were to assume normal interstellar dust, Sirius would have been three magnitudes fainter than it appears today. it seems unlikely that Sirius would have attracted any particular attention as a bright star. In contrast, Schlosser and Bergmann find Babylonian sources that mention that Sirius was visible in the day time and must have been brighter than it is today. We must therefore find a more attractive explanation for the historical observations.

In recent years our picture of white-dwarf evolution has changed quite dramatically. There is now evidence that residual hydrogen burning is still occurring in hot-hydrogen-atmosphere (DA) white dwarfs like Sirius B. Recent theoretical efforts have demonstrated that nuclear shell burning, either by proton–proton reactions or by the diffusion-induced ^{12}C(p, γ)^{13}N can arise[11–13].

Burning by the ^{12}C(p, γ)^{13}N reaction is expected first to arise in DA white dwarfs like Sirius B ($T_{\rm eff} = 27{,}000$ K, where $T_{\rm eff}$ is the stellar effective temperature) once they have cooled to temperatures less than about 30,000 K (refs 11, 12). This hydrogen-burning occurs as the diffusion tail of the hydrogen layer extends downward into the underlying helium layer, where it encounters the diffusion tail of carbon extending upward from the ^{12}C–^{16}O core.

The great sensitivity of the CN reactions to temperature could lead to a thermonuclear shell flash. However, the thermal stability of diffusion-induced reactions in a 1.05 $M_\odot$ white dwarf like Sirius B has not yet been explored either with evolutionary-model calculations or by a linearized analysis of thermal stability. Here we extend the results of Iben and Tutukov[13] and Michaud and Fontaine[12], who discuss white dwarfs of 0.6 $M_\odot$ with ^{12}C–^{16}O cores, to the more massive Sirius B.

The model we present is that Sirius B did become a white dwarf some 10^7 years ago, as is generally accepted. However, it has cooled enough ($T_{\rm eff} < 30{,}000$ K) to initiate a thermonuclear runaway via the ^{12}C(p, γ)^{13}N reaction. It has been shown[14,15] that along its constant-radius cooling track where $\partial L/\partial M < 0$, a steadily burning white dwarf would be thermally unstable. In fact the steady-burning solutions show a high-luminosity stable branch and a low-luminosity unstable branch separated at the point of maximum $T_{\rm eff}$ (bifurcation point) in the Hertzsprung–Russell diagram or in the $M_{\rm e}$–$T_{\rm eff}$ plane, where $M_{\rm e}$ is the mass of hydrogen above the burning region (see Figs 2 and 3 in ref. 16).

It is well known that the long-term behaviour of evolutionary models of actual white dwarfs, even with the steady-burning assumption removed, is still morphologically similar to that of the solutions of the steady-burning model shown in Fig. 3 of Iben[16,17]. In other words, the evolution in the $M_{\rm e}$–$T_{\rm eff}$ plane as described in that figure is valid not only for accretion-induced thermal runaways, but also for a non-accreting star (or one with negligible hydrogen accretion compared with hydrogen lost by burning) experiencing a thermal runaway. The main effect of the hydrogen-burning runaway is to provide the energy necessary to allow the 'thin' hydrogen envelope to make a phase transition across a stability barrier from an unstable to a stable branch with very little change in the mass of the envelope.

Based on the above, we can describe the evolution of Sirius B during the proposed runaway event by applying the solutions or 'tracks' given by Figs 2 and 3 in Iben[16] for a 1.1 $M_\odot$ white dwarf. Therefore, if Sirius B did indeed undergo a thermal pulse, it then expanded to a giant structure, arriving on the stable branch, at a point determined by the mass of hydrogen lying above the burning region ($M_{\rm e}$). If, as we assume, $M_{\rm e} \simeq 10^{-4}$ $M_\odot$, Sirius B would have spent ~250 years mimicking a red giant, burning its envelope hydrogen stably along the plateau (stable) branch to the bifurcation point near maximum $T_{\rm eff}$. From that point, it would proceed to cool rapidly, on a thermal timescale, downward to lower luminosity along the unstable branch. For larger $M_{\rm e}$, longer times are spent on the stable branch. The expansion of Sirius B to a giant structure may have been accompanied by some mass loss but the higher thermal conductivity of a massive white dwarf such as Sirius B could have weakened the flash by the efficient conduction of energy below the burning region. Indeed, the upper limit (3.6×10^{-5} $M_\odot$) for the mass of the ejecta about Sirius B tends to support this interpretation.

The qualitative model presented here is speculative, but considering the growing evidence of the importance of nuclear burning in white dwarfs, plus the physical and observational constraints placed upon the white-dwarf cooling times, models of thermal runaways in white dwarfs offer the best means of describing the historical data for Sirius B. We cannot dispose of the possibility of a thermal runaway triggered by accretion, but the current environment of Sirius B implies low accretion rates both from Sirius A and the interstellar medium. In any event, the historical records may offer a unique observational opportunity to test current theoretical models for white-dwarf evolution.

Note added in proof. An article by H. Takeuchi *et al.* in *Newton* **4**, 138 (1984), mentions the Babylonian record of 700 BC, in which Sirius was observed as a red star. They consider the possibility that Sirius B was a red giant in 700 BC.

but dismiss the idea because a red giant would not have evolved fast enough to become a white dwarf in only 2000 years They speculate that Sirius B was a red giant tens of thousand years ago

Received 27 May, accepted 26 August 1986

1 Schlosser, W & Bergmann, W *Nature* **318**, 45-46 (1985)
2 Böhm-Vitense, E, Dettmann, T & Kapranidis, S *Astrophys J* **232**, L189-L192 (1979)
3 Bruhweiler, F & Sion, E *Astrophys J* **304**, L74 (1986)
4 Bruhweiler, F & Kondo, Y *Astrophys J* **259**, 232-243 (1982)
5 McClintock, W, Henry, R, Linsky, J L & Moos, H W *Astrophys J* **225**, 465-481 (1978)
6 Kondo, Y, Talent, D, Barker, E, Dufour, R & Modisette, J *Astrophys J* **220**, L97-L102 (1977)
7 Aumann, H H *Bull. Am astr Soc* **16**, 483 (1984)
8 Aumann, H H *et al. Astrophys J* **278**, L23-L27 (1984)
9 Smith, B & Terrile, R *Science* **226**, 1421-1424 (1984)
10 Kondo, Y & Bruhweiler, F *Astrophys J* **291**, L1-L5 (1985)
11 Michaud, G, Fontaine, G & Charland, Y *Astrophys J* **280**, 247-254 (1984)
12 Michaud, G & Fontaine, G *Astrophys J* **283**, 787-793 (1984)
13 Iben, I & Tutukov, M *Astrophys J* **282**, 615-630 (1984)
14 Sienkiewicz, R. *Astr Astrophys* **85**, 295-301 (1980)
15 Papaloizou, J, Pringle, J & MacDonald, J *Mon. Not R. astr Soc* **198**, 215-220 (1982)
16 Iben, I *Astrophys J* **259**, 244-266 (1982)
17 Sion, E & Starrfield, S *Astrophys J* **303**, 130-135 (1986)

Ice core record of the $^{13}C/^{12}C$ ratio of atmospheric CO_2 in the past two centuries

H. Friedli, H. Lotscher, H. Oeschger, U. Siegenthaler & B. Stauffer

Physics Institute, University of Bern, Sidlerstrasse 5,
CH-3012 Bern, Switzerland

The release of carbon into the atmosphere due to the activities of humans has caused an increase in concentration as well as a change in the isotopic composition of atmospheric carbon dioxide. CO_2 derived from fossil fuel combustion and from biomass destruction have $\delta^{13}C$ values of ~−25‰ (compared to the atmospheric value of ~−7‰) and are thus depleted in ^{13}C. We have measured $\delta^{13}C$ of CO_2 separated from air trapped in bubbles in ice samples from an ice core taken at Siple Station in Antarctica, in which it has been possible to demonstrate the atmospheric increase of CO_2 (ref. 1) and methane[2] with high time resolution. The isotopic results, together with the CO_2 record from the same ice core, yield information on the sources of excess carbon dioxide and provide a data base for testing the consistency of global carbon cycle models.

Prerequisites for obtaining a good ice core record of the recent past are a high accumulation rate which yields good time resolution, and the absence of summer melting (meltwater interferes with CO_2) Both conditions are fulfilled at Siple Station (75° 55′ S, 83° 55′ W) where the mean annual air temperature is −24 °C and the mean annual accumulation rate is 500 kg m^{-2} The ice core which we studied was cut during the Antarctic summer of 1983-84 to a depth of 200 m by the Polar Ice Coring Office (Nebraska) and the Physics Institute of the University of Bern Enclosure and age of the air have been discussed elsewhere[1,3] It has been possible to date the ice with an accuracy of ~±2 years down to a depth of 144 m (which corresponds to the year 1834) by counting seasonal variations in electrical conductivity[4] Below that depth, the core has been dated by extrapolation

The enclosed air is younger than the surrounding ice, because the enclosure of air in bubbles occurs only between depths of 64-76 m At shallow depths atmospheric air still circulates through the open pores In agreement with earlier papers[1,2], we adopted a value of 83 years for the age difference, and 22 yr for the width of the age distribution (time interval from 10-90% close-off of the total bubble volume)

The experimental procedure has been described previously[5,6] The air was extracted by grinding the ice samples (each ~700 g) at −20 °C in an evacuated container with a milling cutter The air escaping from the mechanically opened bubbles (~80 ml STP per sample) was collected by condensation at 14 K in a small steel cylinder Extraction efficiency is 80-90%, because not all bubbles are opened A small amount of the evolved gas was used for gas-chromatographic measurement of the CO_2 concentration (accuracy ~±3 p p m) CO_2 was quantitatively separated from the remainder of the sample by freezing it out at liquid nitrogen temperature in a metal vacuum line Isotope ratios were measured by mass spectrometry, and the results were corrected for the presence of N_2O which has the same isotopic masses (44, 45 and 46) as CO_2, following the procedure described by Mook and van der Hoek[7], which we checked and confirmed by analyzing artificial N_2O/CO_2 mixtures For $\delta^{13}C$ corrections, the N_2O concentration in the CO_2 sample was determined by mass-spectrometry from the abundance of mass 30, which corresponds to the fragment NO^+ produced in the ion source Details of this procedure will be reported in a separate paper We observed variable and significantly higher N_2O concentrations than expected Tests run afterwards showed that N_2O was produced in the vacuum line in a cold cathode manometer Accuracy of the $\delta^{13}C$ results is 0 10‰, as determined by analysing artificial CO_2-in-air-mixtures and from the scatter of the Siple results around a smooth line The accuracy of the calibration, as determined from analysing international isotope standards, is better than ±0 10‰

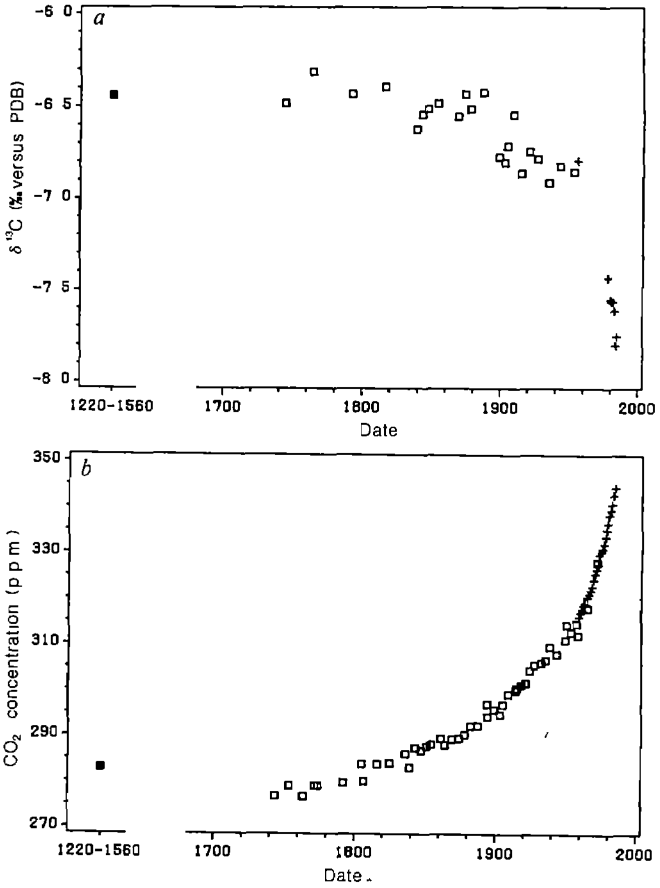


Fig 1 *a*, $\delta^{13}C$ of CO_2 (in ‰ versus PDB standard[18] N_2O-corrected), *b*, CO_2 concentration in air extracted from ice cores from Siple Station (open squares) (data from this work and ref 1) and from South Pole Station (black square)[6] Also shown (crosses) are results from direct atmospheric samples at Mauna Loa (refs 8-10 and C D Keeling, personal communication)

The $\delta^{13}C$ value and the concentration of CO_2 in the Siple ice samples are presented in Table 1 and Fig 1 For comparison, the average values for air extracted from an interval of a South Pole ice core which is between 430 and 770 years old[6], and results from direct atmospheric measurements at Mauna Loa (refs 8-10 and C D Keeling, personal communication) are also

Table 1 Analytical results of the Siple ice core

Average depth (m)	Gas age (yr AD)	Concentration (p p m)	$\delta^{13}C$ (‰, PDB)
187 70	1744	276 8	−6 48
177 50	1764	276 7	−6 31
168 30	1791	279 7	−6 43
154 89	1816	283 8	−6 39
142 75	1839	283 1	−6 62
140 75	1843	287 4	−6 54
138 20	1847	286 8	−6 51
134 47	1854	288 2	−6 48
126 80	1869	289 3	−6 55
123 80	1874	289 5	−6 43
121 80	1878	290 3	−6 51
116 82	1887	292 3	−6 42
110 20	1899	295 8	−6 77
108 80	1903	294 8	−6 80
107 20	1905	296 9	−6 71
105 25	1909	299 2	−6 54
101 80	1915	300 5	−6 86
98 80	1921	301 6	−6 74
95 17	1927	305 5	−6 78
90 77	1935	306 6	−6 91
86 80	1943	307 9	−6 82
81 22	1953	312 7	−6 85

shown in Fig 1 The average $\delta^{13}C$ value of the three samples from before 1800 (−6 41‰) is in excellent agreement with the mean value for the South Pole samples (−6 44 ± 0 03‰) This strengthens our confidence in the reliability of the experimental method We cannot exclude the possibility that the isotopic composition of CO_2 trapped in ice is shifted compared to the atmospheric value However, the agreement indicates that this is not the case, because isotopic fractionation processes active during trapping or storage of the air bubbles in ice could be expected to have different effects on the two cores which sample ice accumulated at different temperatures (the mean annual temperature is −24 °C at Siple Station, and −51 °C at the South Pole)

During the nineteenth and twentieth centuries, $\delta^{13}C$ slowly decreased, approximately antiparallel to the CO_2 concentration trend The youngest ice core data fit well the direct atmospheric observation in 1956 of −6 79‰ (results given in ref 8 and N_2O-corrected according to Mook and van der Hoek[7]) The overall decrease from the end of the eighteenth century to 1980 (−7 55‰ at Mauna Loa[9]) is −1 14‰ The estimated overall uncertainty of this figure is ±0 15‰, including a possible difference in calibration between the two data sets

The $\delta^{13}C$ record from the Siple ice core exhibits a step-like shift before 1900 and a plateau between ~1920 and 1950 There is also a less well pronounced plateau in the CO_2 results at the same depth It is premature at this stage to decide whether this actually reflects an accelerated atmospheric CO_2 and $\delta^{13}C$ change around the turn of the twentieth century, or whether it represents an artefact

The ice-core results can be compared with tree-ring $\delta^{13}C$ records, measured by a number of authors[11,12] Tree-ring $\delta^{13}C$ values generally exhibit considerable scatter, due to the fact that the $^{13}C/^{12}C$ fractionation varies from plant to plant as well as from year-ring to year-ring within a plant Different records, each an average over a number of trees, have yielded changes from 1800 to 1980 ranging from 0‰ (ref 11) to −2‰ (ref 12) Thus, it is difficult to compare those results with ours Stuiver *et al*[13] attempted to correct for variable physiological isotope fractionation by making use of a correlation between $\delta^{13}C$ and tree-ring area, and obtained a change of −0 8‰ from 1800 to 1980 Interestingly, many trees show constant (or even increasing) values in the period 1920 to 1950[12,14], which is also indicated by our results

The CO_2 results (Fig 1*b*) include, besides the IR laser absorption data of Neftel *et al*[1], additional results obtained by gas chromatography They indicate a pre-industrial (pre-1800) concentration of ~280 p p m , which is nearly equal, within a few p p m , to the average for the South Pole samples (283 ± 5 p p m , age interval 430–770 years) We therefore have no reason to assume any secular fluctuation exceeding 10 p p m between the years of 1200–1800

In order to compare the $\delta^{13}C$ and the CO_2 concentration results and to interpret these results in terms of CO_2 release rates into the atmosphere, it was necessary to use a model of the global carbon cycle[15,16] Within the framework of the carbon cycle model, the CO_2 and the $\delta^{13}C$ records from the Siple ice core are consistent, the calculated $\delta^{13}C$ trend (using the CO_2 concentration history of Fig 1*b* as model input) agrees well with the observations shown in Fig 1*a* Taking the Siple CO_2 and $\delta^{13}C$ records as model inputs, we calculated the release rates of CO_2 into the atmosphere The results depend on the structure of the oceanic and biospheric sub-models, but indicate that the cumulative non-fossil CO_2 production until 1980 was about half of the cumulative fossil production, and 50% or more of this non-fossil CO_2 was released prior to 1900

What additional information can be obtained from isotopic data that is not already contained in the CO_2 concentration record? Fossil and biospheric carbon both have very similar $\delta^{13}C$ values, but oceanic carbon is considerably richer in ^{13}C Thus, $\delta^{13}C$ can, in principle, be used to decide whether the non-fossil CO_2 input, as reconstructed from the concentration record, was of biospheric or of oceanic origin CO_2 escaping from the ocean is isotopically very similar to atmospheric carbon dioxide, so that input from the ocean would not have a significant effect on $\delta^{13}C$ values in contrast to input from the biosphere The difference in atmospheric $\delta^{13}C$ values between a biospheric and an oceanic CO_2 input depends on the input history As will be discussed in detail elsewhere[17], model calculations with the non-fossil input history back-calculated from the CO_2 record of Fig 1*b* yield a relatively small difference between biospheric and oceanic sources in the overall isotopic decrease This result can be understood by considering the different behaviour of CO_2 concentration and isotopic composition with respect to atmosphere–ocean interaction The oceanic uptake capacity for CO_2 is limited compared to that for isotopic perturbations by a chemical buffering mechanism, and an isotopic perturbation is additionally attenuated by exchange with the biosphere Thus, a large CO_2 release during the 19th century still significantly affects the present-day atmospheric CO_2 concentration, but not the $\delta^{13}C$ value In other words, $\delta^{13}C$ values integrate over a relatively short time and therefore, a $\delta^{13}C$ record can reveal the time variations of the CO_2 release rates in greater detail than a record of CO_2 concentrations For this reason, further ice-core measurements are required to establish the past atmospheric history of CO_2 concentration and of $\delta^{13}C$ values

We thank K Hanni for help with mass-spectrometric measurements and CO_2 extraction Supported by the Swiss NSF

Received 23 June, accepted 13 August 1986

1 Neftel, A , Moor, E , Oeschger, H & Stauffer, B *Nature* **315**, 45–47 (1985)
2 Stauffer, B , Fischer, G , Neftel, A & Oeschger, H *Science* **229**, 1386–1388 (1985)
3 Schwander, J & Stauffer, B *Nature* **311**, 45–47 (1984)
4 Schwander, J thesis, Univ Bern (1984)
5 Moor, E & Stauffer, B *J Glaciol* **30**, 358 (1984)
6 Friedli, H , Moor, E , Oeschger, H , Siegenthaler, U & Stauffer, B *Geophys Res Lett* **11**, 1145–1148 (1984)
7 Mook, W G & van der Hoek, S *Isotopic Geosci* **1**, 237–242 (1983)
8 Keeling, C D , Mook, W G & Tans, P P *Nature* **277**, 121–123 (1979)
9 Mook, W G , Koopmans, M , Carter, A F & Keeling, C D *J geophys Res* **88**, 10915–10933 (1983)
10 Heimann, M , Keeling, C D & Mook, W G *Tellus* (in the press)
11 Francey, R J *Nature* **290**, 232–235 (1981)
12 Freyer, H D & Belacy, N *J geophys Res* **88**, 6844–6852 (1983)
13 Stuiver, M , Burk, R L & Quay, P D *J geophys Res* **88**, 11731–11748 (1984)
14 Tans, P P & Mook, W G *Tellus* **32**, 268–283 (1980)
15 Siegenthaler, U , Heimann, M & Oeschger, H in *Carbon Dioxide, Climate and Society* (ed Williams, J) 79–87 (Pergamon, Oxford, 1981)
16 Peng, T H , Broecker, W S , Freyer, H D & Trumbore, S *J geophys Res* **88**, 3609–3620 (1983)
17 Siegenthaler, U & Oeschger, H *Tellus* **39B** (in the press)
18 Craig H *Geochim. cosmochim Acta* **3**, 53–92 (1953)

Seasonal variation of methylarsenic compounds in airborne particulate matter

Hitoshi Mukai*, Yoshinari Ambe*, Tatsunori Mukut†, Kazuo Takeshita† & Tsuneo Fukuma†

* Chemistry and Physics Division, National Institute for Environmental Studies, Yatabe Tsukuba, Ibaraki, 305 Japan
† Saigo Health Center, Minato Saigo, Oki, Simane 685 Japan

Inorganic and organic arsenic compounds in nature are methylated by biological action and released into the air, forming trimethylarsine and dimethylarsine[1–5]. These compounds are oxidized in the air to dimethylarsinic acid and trimethylarsineoxide[6] and exist as particles[7,8]. In this study, we examine seasonal variations of methylarsenic compounds in airborne particulate matter to try to understand the characteristics of the methylation of arsenic in nature. In summer a high concentration of dimethyl and trimethyl forms of arsenic was observed, while in winter the levels were very low. This seasonal variation can be attributed to the different temperatures of the environment during these seasons The activation energy of methylation of arsenic was estimated to be ~12 kcal mol^{-1} from Arrhenius plots and can be used to predict the behaviour of these methylarsenic compounds, which are considered to form an important portion of the arsenic present in airborne particulate matter.

The natural methylation of arsenic has been reported in lakes, seas and soils[9–11] Challenger and Higginbottom[2] reported that bread mould produced trimethylarsine from inorganic arsenic Cox and Alexander[4] also showed that fungi in sewage could produce trimethylarsine from inorganic and methylarsenic compounds Dimethylarsine is synthesized by methanobacterium[3] These arsines are released from various environmental media and are present in air Johnson and Braman[7] detected both gaseous and particulate trimethyl and dimethyl forms of arsenic in the atmosphere The methylation of arsenic is thus important for determining the mobility of this element in nature However, biological methylation in nature is not well understood because the concentrations of these methylated compounds in the air are very low Methylarsenic compounds, for example, are oxidized in the air and converted to particulate form (trimethylarsineoxide and dimethylarsinic acid)[6] Therefore, concentrations of these compounds in airborne particulate matter reflect the activity of biomethylation in nature

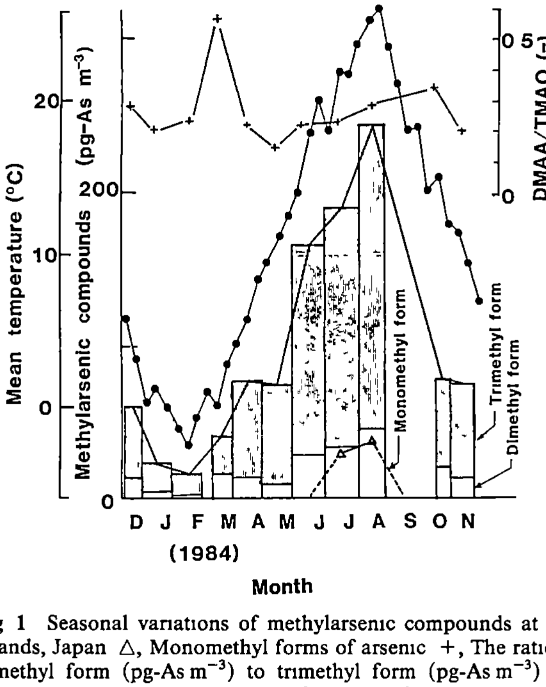


Fig 1 Seasonal variations of methylarsenic compounds at Oki Islands, Japan △, Monomethyl forms of arsenic +, The ratio of dimethyl form (pg-As m^{-3}) to trimethyl form (pg-As m^{-3}) ●, Mean temperature of every 10 days

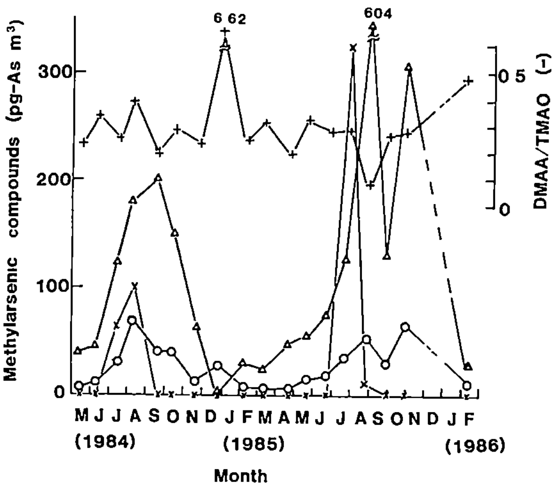


Fig 2 Seasonal variations of methylarsenic compounds at Tsukuba, Japan △, Concentration trimethyl, ○, dimethyl, ×, monomethyl forms of arsenic +, The ratio of dimethyl form to trimethyl form

Sampling of airborne particulate matter was carried out over 1 year at two sites in Japan, Oki Islands (an unpolluted area) in the Japan Sea and Tsukuba (a rural inland area) These two sampling sites are ~1,000 km apart The sampling at Oki Islands was carried out over 1 month with a low volume air sampler at the top of a mountain (200 m) (December 1983–November 1984) The samples from Tsukuba were collected with a high-volume air sampler for 3 days each month (May 1984–February 1986) Quartz fibre filters (PALLFLEX 2500QAST) were used, loaded filters were preserved in a store room at −20 °C

The hydride generation method proposed by Braman *et al*[7,12] was adapted to analyse the methylarsenic compounds in airborne particulate matter Dimethyl and trimethyl forms of arsenic (possibly dimethylarsinic acid and trimethylarsineoxide) were extracted from the filter with 10 cm^3 of 1 M HCl+0 02 M EDTA (9 1) and reduced with 6 cm^3 of 10% $NaBH_4$ in 1M HCl containing 20 μmol EDTA The arsines produced were trapped in a half-packed U-tube dipped in liquid nitrogen Following gas chromatography, each arsine separated was detected as arsenic by atomic absorption spectroscopy The details of this method are described elsewhere[13] The detection limits of monomethyl, dimethyl and trimethyl forms of arsenic by this method were ~2 pg-As m^{-3}

Figure 1 shows the variations of methylarsenic compounds in airborne particulate matter sampled at Oki Islands High concentrations of dimethyl and trimethyl forms were observed in summer The pattern of the seasonal changes of the concentrations of dimenthyl and trimethyl forms was found to correlate with the mean temperature The order of these concentrations was similar to that reported by Johnson and Braman[7] The ratio of the dimethyl form to trimethyl form was 0 15–0 34 in most cases, indicating that trimethylarsine was the dominant form of methylated arsenic produced by biomethylation Traces of the monomethyl form were detected in the July and August samples

Although the ratio of these methylarsenic compounds to inorganic arsenic in airborne particulate matter was <10%, the contribution of methylated arsenic cannot be ignored, because the concentration of methylarsenic compounds in airborne particulate matter often amounted to 18 p p m whereas the concentration of arsenic in the Earth's crust is 2 p p m

Figure 2 illustrates the seasonal variations of methylarsenic compounds at Tsukuba Although the two sampling sites are

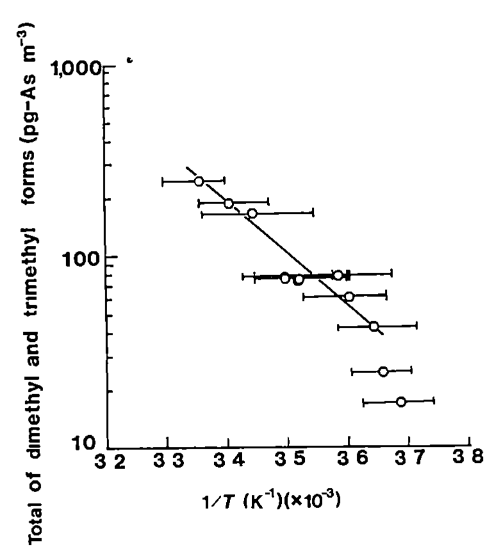


Fig 3 Arrhenius plot using the total concentration of the dimethyl and trimethyl arsenic compounds in the air (particulate form) at Oki Islands The line was drawn using nine points, and disregards two winter points, when snow covered the land surface The horizontal line of each point indicates the range between the mean value of maximum and minimum temperature

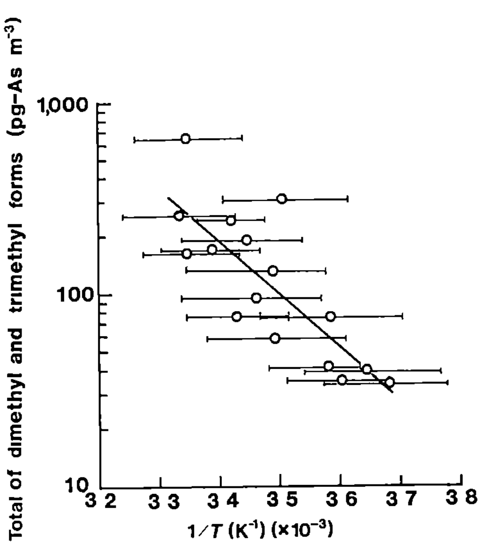


Fig 4 Relation between the total of the concentrations of the dimethyl and trimethyl arsenic compounds (particulate form) at Tsukuba The line obtained from Fig 3 is shown The horizontal line at each point indicates the range between maximum and minimum temperature during the sampling period

far apart from each other, the patterns of the variation and the concentration of the methylarsenic compounds were very similar in Oki Islands and Tuskuba At both locations the concentrations of dimethyl and trimethyl forms were high in summer, but low in winter The monomethyl form, which was detected in the summer samples at Oki Islands, was also detected only in the July and August samples, and was considered to be caused by contamination with a pesticide[14] (iron methane arsonate) spread over rice fields 20-30 km away from the sampling site, not biomethylation Therefore, we shall not consider this form in the following discussion

The seasonal variations of the dimethyl and trimethyl forms of arsenic can be attributed to the change of temperature of the environment Tanaka *et al*[8] reported that a high concentration of dimethylarsenic compound in airborne particulate matter was observed during the summer In the case of methylation of selenium, Chau *et al*[15] pointed that the production of volatile selenium was also temperature dependent Figure 3 shows the relation between the total dimethylarsenic and trimethylarsenic concentrations (mean concentration for ~1 month) at Oki Islands and the reciprocal of the mean temperature of the sampling period If we assume that arsines diffuse into the atmosphere faster than their rate of production, and are oxidized constantly, the concentration of these compounds is considered to reflect the rate of production of methylarsenic compounds, and the activation energy of biomethylation of arsenic in nature is estimated to be 12 kcal mol^{-1} from the slope of the graph (avoiding the two winter samples, using all the points gives a value of ~16 kcal mol^{-1}) This value is consistent with those of general biological reactions (8-18 kcal mol^{-1}) But, as biomethylation is influenced not only by temperature but also by other factors such as humidity[11,16] and the concentration of these compounds in air is affected by meteorological conditions (such as mixing depth, wind speed and weather), the activation energy obtained is considered to be the apparent value It will reflect various environmental factors, even though temperature is a major factor For example, if the rate that the arsine gas is oxidized and converted to the airborne particulate matter is faster in summer, the slope of the line of the Arrhenius plot will be overestimated compared with the real value, whereas the existence of a surface inversion layer in winter might make the concentration of airborne particulate matter relatively higher and lead to an underestimation of the activation energy The effect of an inversion layer may be minor, because the sampling site is located on top of the mountain (200 m) Further study will be necessary in order to understand the conversion rate, which includes the rate of heterogeneous reaction between arsine gas and airborne particulate matter as well as the rate of homogeneous oxidation reaction with oxygen[6], peroxides and hydroxyl radicals

Figure 4 shows the data from Tsukuba Many points were distributed along the line drawn from the data in Fig 3, indicating that a similar activation energy can be obtained even though the sampling sites are different The deviation of two points from the line suggests that the data obtained during short-period sampling is strongly influenced by the meteorological conditions at the time and that the line does not describe the maximum concentration corresponding to each temperature

Our data showed that the methylation of arsenic in nature is an important source of arsenic in the air and is influenced by temperature The activation energy estimated above can be used to predict the behaviour of methylarsenic compounds in the air throughout the year, although it is not known to which methylation reaction this value corresponds These findings do not give any new insight into biological methylation, however, they give important information concerning the contribution of biomethylation to the geological cycles of trace elements (As, Se and so on) The data shown here were only from the particulate form, the gaseous form of trimethylarsine needs further investigation The measurements of the change of the rate of the production of methylarsenic compounds with the temperature of variables such as the land surface, which is one of the most probable sources of methylarsenic compounds, are also needed for comparison with the obtained activation energy

We thank members of Saigo Health Center and NTT (Saigo) for the field assistance, Dr T Uehiro, Dr M Morita and Dr Y Yokouchi for useful advice

Received 9 June, accepted 3 September 1986

1 Challenger, F, Higginbottom, C & Ellis, L *J chem Soc* 95-101 (1933)
2 Challenger, F & Higginbottom, C *Biochem J* **29**, 1757-1788 (1935)
3 McBride, B C & Wolfe, R S *Biochemistry* **10**, 4312-4317 (1971)
4 Cox, D P & Alexander, M *Bull Envir Contam Toxicol* **9**, 84-88 (1973)
5 Wood, J M *Science* **183**, 1049-1052 (1974)
6 Parris, G E & Brinckman, F E *Envir Sci Technol* **10**, 1128-1134 (1976)

7 Johnson, D L & Braman, R S *Chemosphere* **4**, 333-338 (1975)
8 Tanaka, S, Kaneko, M & Hashimoto, Y *J chem Soc Jap* 637-642 (1984)
9 Braman R S & Foreback, C C *Science* **182**, 1247-1249 (1973)
10 Andreae, M O *Limnol Oceanogr* **24**, 440-452 (1979)
11 Akins, M B & Lewis, R J *Soil Sci Soc Am J* **40**, 655-658 (1976)
12 Braman, R S, Johnson, D L, Foreback, C C, Ammons, J M & Bricker, J L *Analyt Chem* **49**, 621-625 (1977)
13 Mukai, H & Ambe, Y *Analyt chim Acta* (submitted)
14 Mukai, H & Ambe, Y *Atmos Envir* (in the press)
15 Chau, Y K, Wong, P T S, Silverberg, B A, Luxon, P L & Bengert, G A *Science* **192**, 1130-1131 (1976)
16 Woolson E A *Envir Hlth Perspect* **19**, 73-81 (1977)

Reversed magnetization in pyroclastics from the 1985 eruption of Nevado del Ruiz, Colombia

Friedrich Heller*, Juan Carlos Carracedo†
& Vicente Soler†

* Institut fur Geophysik, ETH-Honggerberg, CH-8093 Zurich, Switzerland
† Estacion Volcanologica de Canarias,
Instituto de Recursos Naturales, (CSIC), La Laguna,
Tenerife, Canary Islands, Spain

One of the fundamental principles of palaeomagnetism and rock magnetism states that the direction of natural remanent magnetization (NRM) is acquired parallel to the applied field. However, ferromagnetic mineral phases with complex microstructures may violate this rule[1,2] and produce a magnetization vector which is antiparallel to the direction of the external field. Andesitic pumice, which was hurled several kilometres during the disastrous 1985 eruption of the Nevado del Ruiz volcano (Colombia), carries a stable but reversed NRM with southerly declination and negative inclination. Heating experiments show that this magnetization is due to a self-reversal mechanism which also induces a reversed thermoremanent magnetization (TRM) in the laboratory field. Thus we have used the Nevado del Ruiz 1985 pumice, which was magnetized in the present well-known geomagnetic field, to demonstrate for the first time that self reversal has actually controlled the process of NRM acquisition.

The Nevado del Ruiz, a stratovolcano of the Ruiz-Tolima complex in the northernmost part of the Andean Cordillera (lat 4°53′ N, long 75°10′ W, altitude 5,400 m, regional geomagnetic field inclination +31°), covered by a 17 km^2 ice cap, has had two major eruptions in 1595 and 1845, and its last explosive eruption was on 13 November 1985 (ref 3) Lithic fragments, ash, and juvenile scoria and pumice fragments were emitted during the last eruption and an eruptive column up to $\sim$5,600 m in height was formed Most of the hot juvenile eruptive products fell on the ice cap The catastrophic mudflow thus generated travelled down the radial river valleys causing the death of 25,000 people and damage assessed at US$ 300 million

Dacitic andesite pumice fragments were spread around the emission centre in the form of pyroclastic falls Some of the fragments were embedded in the mud as they fell, or penetrated cushions of the *Plantago rigida* vegetation Where they lodged, moderate thermal effects were observed in smashed plant fragments

To investigate the temperature of deposition of this pumice, hand samples of encrusted fragments (5-10 cm in diameter), oriented with a magnetic compass, were collected within a radius of 5-6 km from the emission centre The oriented samples were collected over a 3-week period after the 13 November eruption by one of us (J C C), who formed part of the Scientific Commission sent by the Spanish Ministry of Foreign Affairs to collaborate in the study and surveillance of the Nevado del Ruiz volcano

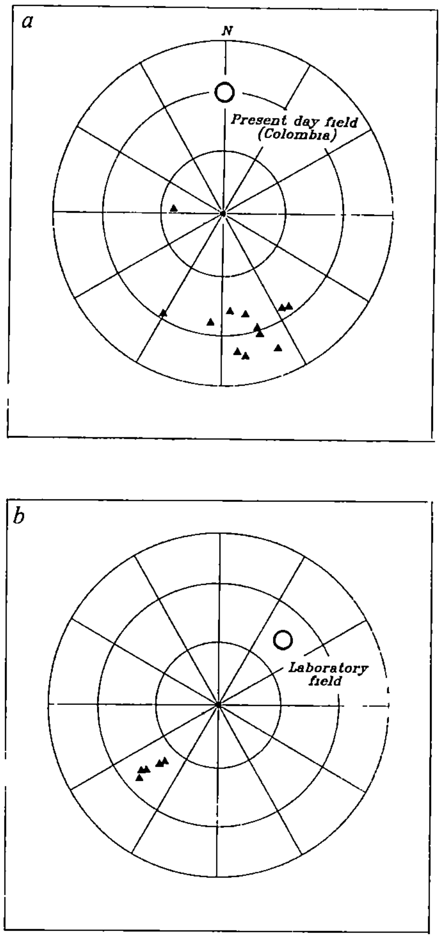


Fig. 1 Directions of *a*, natural remanent magnetization (NRM), mean direction declination = 170 7°, inclination = −35 9°, α_{95} = 14 2, 12 samples plotted *b*, Laboratory-induced thermoremanent magnetization (TRM), mean direction declination = 226 6°, inclination = −44 7°, α_{95} = 6 3, 5 samples plotted Plots are stereographic projections together with applied field directions Field and magnetization vectors are antiparallel ○, Lower, ▲, upper hemisphere

The result of the initial magnetic remanence measurements is shown in Fig 1 All samples carry a reversed NRM which—within statistical error—is antiparallel to the present field of the region Alternating field (AF) demagnetization reduces the directional scatter that results from small orientation inaccuracies and from additional magnetization components of normal polarity and of spurious direction The main reversed NRM component is stable up to 80 mT and has median destructive fields between 10 mT and 30 mT

When the NRM vector is measured during repeated heating and cooling cycles at variable temperatures in air and in zero field[4], a polarity change is registered at temperatures $\sim$120 °C (Fig 2), the strong reversed room temperature component being completely removed at 150-160 °C Above these temperatures the NRM polarity is always normal (parallel to the present field) When cooling back to room temperature after the first heating cycle, the reversed component develops again in several samples (Fig 2*a*) Thus, the reversal mechanism is still active in these laboratory conditions On heating to higher temperatures the normal NRM component disappears between 350 and 400 °C

Two optically homogeneous magnetic mineral phases of different colour and reflectivity are observed using the ore microscope A brownish, low reflectivity but strongly magnetic phase is attributed to titanomagnetite A whitish-yellow, higher reflec-

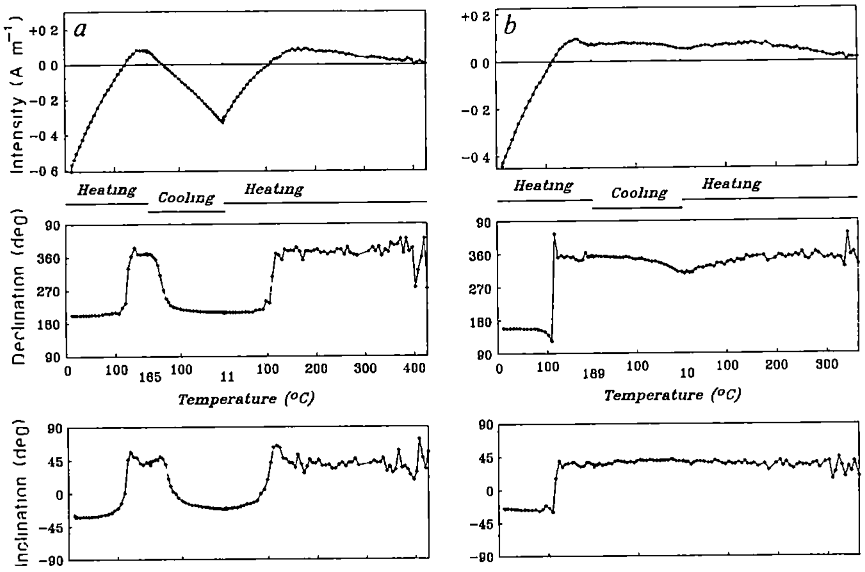


Fig. 2 Inclination, declination and intensity of natural remanent magnetization of two pyroclastic samples during thermal treatment in air and zero magnetic field Heating rate 10 °C min^{-1} *a*, The NRM direction reversibly changes polarity and NRM intensity changes sign at ~120 °C on first heating to 165 °C, during subsequent cooling to room temperature and in the second heating cycle up to 425 °C This indicates that a self-reversal process was active during initial NRM formation *b*, After removal of the reversed component during the first heating cycle up to 189 °C, only a normal NRM component is observed in the subsequent cooling and heating cycles Note that the temperature scale has been folded around the maximum temperature of the first heating cycle and the minimum temperature of the subsequent cooling cycle

tivity mineral which is less magnetic but mechanically very hard and is always closely associated with the titanomagnetite, is identified as ilmenohaematite An appreciable titanium content in these oxides is suggested by the intermediate Curie temperatures which were found to be between 370 and 410 °C for the major component of the saturation magnetization and between 180 and 200 °C for the minor component

Both magnetic phases are extraordinarily resistant to oxidation during thermal treatment in air, as indicated by complete reversibility of strong field magnetization behaviour up to temperatures >600 °C Therefore, a TRM was given to some samples between 450 °C and room temperature in the presence of a laboratory field of the same magnitude as the present geomagnetic field in Colombia (Fig 1*b*) The TRM vector acquired again has opposite polarity to the applied field and this must be due to self reversal AF cleaning of this TRM shows that a normal component with low coercivity and small magnitude (up to 30% of the reversed component) is superimposed NRM- and TRM-intensities are equal after AF demagnetization at a field strength of 10 mT and demonstrate that the NRM was acquired as a TRM over the temperature range mentioned In accordance with observations reported from the explosion[3], this points to a hot emplacement of the pumice bombs at temperatures higher than the maximum Curie temperatures

The evidence that oppositely directed NRM directions are caused by frequent polarity reversals of the geomagnetic field is overwhelming Nevertheless, the occurrence of self reversal of NRM in igneous rocks may be underestimated To qualify NRM directions as being parallel or antiparallel to the magnetizing field, much more detailed and time consuming experiments have to be conducted than those which are usually performed during routine palaeomagnetic measurements Therefore, a substantial amount of self reversals may have been overlooked in the past Possibly some of the abnormal NRM polarities in Mankinen and Dalrymple's[5] polarity record for the past 5 Myr may be due to self-reversing basalts

The physical models of self reversal can be divided into two categories single-phase and double-phase models The sign of spontaneous magnetization in Néel's[6] N-type ferrites may change at a specific temperature due to different temperature dependence of the sublattice magnetizations Negative magnetic coupling between two phases may result from magnetostatic or exchange interaction All these mechanisms have been observed to act on the TRM of certain igneous rocks[7-11]

The self-reversal mechanism of the Nevado del Ruiz 1985 pumice is not yet fully understood The one-phase model, however, can be excluded because the temperature dependence of saturation magnetization shows normal behaviour A negative coupling between the two visible ore phases seems to be unlikely because the continuous demagnetization experiments (Fig 2) show the reversibility of the self-reversal effect in laboratory conditions to be very sensitive to the maximum demagnetization temperatures between 160 and 190 °C The higher the temperature is raised in this temperature interval, the more suppressed is the self-reversal effect Possibly two intimately mixed, chemically very similar phases within the lower Curie-point phase are interacting If a high enough temperature is reached, the normally magnetized phase with slightly higher Curie temperature becomes demagnetized so that on cooling in zero field no interaction takes place with the other phase of slightly lower Curie temperature Both phases remain demagnetized on recooling to room temperature and the remaining remanent magnetization originating from the high-Curie-point phase has normal polarity (Fig 2*b*) Such a mechanism would closely resemble the exchange interaction model developed by Uyeda[10]

The major significance of the self-reversal observed in the Nevado del Ruiz 1985 pumice, however, lies in the clear evidence that this mineralogical mechanism has indeed controlled the polarity of the NRM when the rock formed

We thank Professors J M Fúster and A Cendrero and Drs J C Thouret and H Cepeda for help with collecting samples We benefited from conversations about the Nevado del Ruiz 1985 explosion with Dr B Martinelli and received help with measurements from W Gruber Professor W Lowrie's suggestions improved the manuscript Contribution 523, Institut für Geophysik, ETH Zurich

Received 9 July, accepted 18 September 1986

1 Stacey, F D & Banerjee, S K in *The Physical Principles of Rock Magnetism*, 1-195 (Elsevier, Amsterdam, 1974)
2 O'Reilly, W in *Rock and Mineral Magnetism*, 1-220 (Blackie, Glasgow, 1984)
3 Herd, D G & Comite de Estudios Volcanológicos *Eos* **67**, 457-460 (1986)
4 Heiniger, C & Heller, F *Geophys J R astr Soc* **44**, 281-288 (1976)
5 Mankinen E A & Dalrymple, G B *J geophys Res* **84**, 615-626 (1979)
6 Neel L *Adv Phys* **4**, 191-243 (1950)
7 Schult, A *J Geophys* **42**, 81-84 (1976)
8 Ozima, M & Ozima, M *Earth planet Sci Lett* **3**, 213-215 (1967)
9 Sasajima, S & Nishida, J *Rock Magn Paleogeophys* **2**, 5-9 (1974)
10 Uyeda, S *Jap J Geophys* **2**, 1-123 (1958)
11 Heller, F & Petersen, N *Phys Earth planet Inter* **30**, 358-372 (1982)

Geochemical correlation between southern African kimberlites and South Atlantic hotspots

Anton P. le Roex

Department of Geology, University of Cape Town, Rondebosch,
7700 South Africa

Recent isotopic and trace element data[1,2] have led to the subdivision of southern African Cretaceous/Jurassic kimberlites into group I and group II varieties. The former have Sr, Nd and Pb isotopic signatures similar to those commonly associated with ocean island basalts and are thought to derive from within the asthenosphere[1]. The high Sr and low Nd and $^{206}Pb/^{204}Pb$ isotopic ratios of group II kimberlites have been cited as evidence that these kimberlites are derived from a source within the sub-continental lithosphere[1,3]. Basalts erupted at the surface expression of South Atlantic hotspots preserve isotopic and trace element ratios similar to those of group I and group II kimberlites. Hotspots falling within the Dupal anomaly[4] have similar characteristics to group II kimberlites and those outside the Dupal anomaly to group I kimberlites. This correlation raises the possibility that the source region of southern African group II kimberlites is also located within the asthenosphere (within recycled lithospheric material) and not necessarily within the sub-continental lithosphere.

Southern African Cretaceous/Jurassic kimberlites comprise two dominant petrographic and geochemical varieties, namely so-called micaceous (group II) and basaltic or non-micaceous (group I) kimberlites[1,5]. Both types contain diamonds, appear to be distributed throughout the Kaapvaal craton and show a broad age correlation with group I kimberlites (80–114 Myr), being generally younger than group II kimberlites (114–200 Myr)[1,6]. Group I kimberlites have significantly lower $^{87}Sr/^{86}Sr$ and Rb/Sr ratios and higher $^{143}Nd/^{144}Nd$, $^{206}Pb/^{204}Pb$ and U/Pb ratios than group II kimberlites[1-3,5,7].

Smith[1] has argued that these geochemical differences are not a result of alteration and/or crustal contamination but reflect substantial chemical differences in their mantle source regions. The slightly depleted source region signature of group I kimberlites, relative to bulk Earth, has been cited as evidence for an asthenospheric origin (their isotopic characteristics are similar to most ocean island basalts). In contrast, the enriched source region signature of group II kimberlites has been used to argue for an origin within the ancient enriched sub-continental lithosphere[1-3].

Hart[4] has recently delineated a major region between southern Africa and South America that forms part of his proposed Dupal anomaly, in which the isotopic signature of hotspot lavas is distinct from those occurring to the north of ~20° S and to the south of ~43° S. More recent data (S. R. Hart, unpublished data) shows that in the South Atlantic this anomaly extends as far south as the recently documented Shona hotspot at ~54° S[8,9]. Basalts that have erupted at hotspots within the Dupal region have Sr and Nd isotopic ratios enriched compared with the bulk Earth and have higher $^{207}Pb/^{204}Pb$ and $^{208}Pb/^{204}Pb$ ratios than most other oceanic volcanics. Here I compare the isotopic and trace element 'signatures' of South Atlantic hotspots and southern African kimberlites.

Basalts from the South Atlantic describe two distinct trends in terms of Zr/Nb ratio versus $^{87}Sr/^{86}Sr$ and $^{143}Nd/^{144}Nd$ ratios (Fig. 1). The more radiogenic Sr and less radiogenic Nd trends for a given Zr/Nb ratio correspond to basalts erupted within the Dupal anomaly and comprise the hotspots of Gough, Discovery and Tristan da Cunha and MORB in the vicinity of the Shona hotspot and, by implication, the as yet unsampled Shona hotspot[8]. The lower trend in each diagram includes the hotspots of Bouvet, Marion and Ascension and MORB from the vicinity of the Bouvet hotspot. It is evident from Fig. 1 that group I kimberlites plot on the extension of the 'Bouvet' trend whereas group II kimberlites plot on the extension of the Dupal trend.

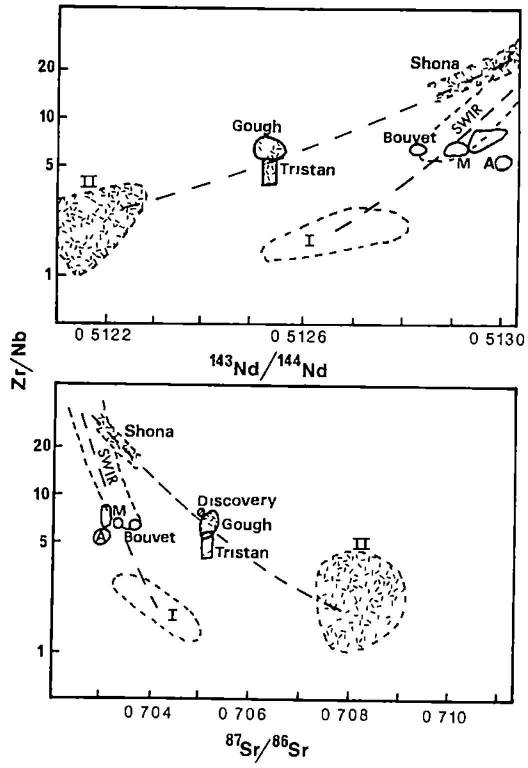


Fig. 1 Correlation between Zr/Nb ratio and $^{87}Sr/^{86}Sr$ and $^{143}Nd/^{144}Nd$ ratios in South Atlantic basalts and group I and group II kimberlites. Initial isotope ratios of the kimberlites have been used for plotting. SWIR, Southwest Indian Ridge, II, group II kimberlites, I, group I kimberlites, A, Ascension, M, Marion (two fields). Sources of data refs 1, 3, 8, 19, 25–28, 31, 34, S. R. Hart (unpublished data) and A. P. le R. (unpublished data).

A similar correlation exists for incompatible trace element variations. Figure 2 shows the variation of Ba/Nb, Zr/Nb and La/Nb ratios in MORB, South Atlantic hotspot basalts and group I and group II kimberlites. Hotspots within the Dupal anomaly (that is, Gough, Discovery, Tristan da Cunha and the associated Walvis Ridge and Rio Grande Rise) are characterized by high Ba/Nb (8–22), La/Nb (0.75–1.25) and low Zr/Nb (<10) ratios and define a trend extending away from the field of Bouvet, Marion and other non-Dupal-area hotspots towards group II kimberlites (Ba/Nb > 15, La/Nb > 1, Zr/Nb < 3). Although data are not yet available from the Shona hotspot, the high Ba/Nb ratios of associated southern Mid-Atlantic Ridge enriched MORB (Fig. 2) indicates that this hotspot is also likely to be characterized by high Ba/Nb ratios. Group I kimberlites, on the other hand, have low Ba/Nb (<8), La/Nb (<0.8) and Zr/Nb (<4) ratios, similar to those of the non-Dupal hotspots Bouvet, Marion and Ascension (Ba/Nb < 9, La/Nb < 0.6 and Zr/Nb < 10). Depleted MORB in these diagrams defines a very distinct trend extending towards high Zr/Nb (>20) and La/Nb (>1.0) ratios at low (<7) Ba/Nb ratios.

Figure 3 extends this correlation between group II kimberlites and South Atlantic Dupal hotspot volcanics to include Pb isotopes. It is evident that these lavas all plot on a sub-parallel Pb isotope array characterized by low $^{206}Pb/^{204}Pb$ and high $^{208}Pb/^{204}Pb$ and (excluding Tristan da Cunha) $^{207}Pb/^{204}Pb$ ratios relative to MORB, non-Dupal-area hotspot volcanics and group I kimberlites. The high ^{208}Pb and ^{207}Pb of enriched southern Mid-Atlantic Ridge basalts (Fig. 3, S. R. Hart, unpublished data) associated with the Shona hotspot indicate that the latter is likely to have Pb isotopic signature similar to that of Gough

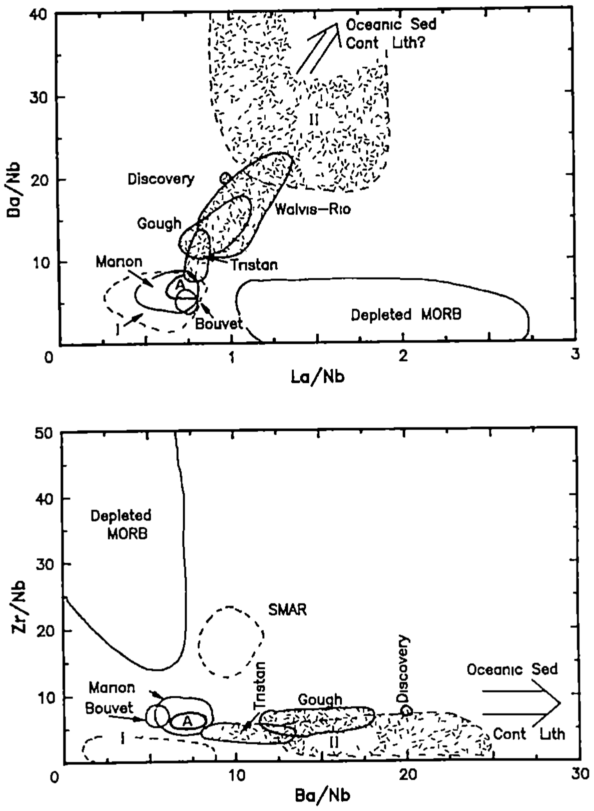


Fig 2 Covariation of Ba/Nb, Zr/Nb and La/Nb ratios in South Atlantic basalts and group I and group II kimberlites Compositions of oceanic sediment[19] and sub-continental lithosphere[26] are shown for comparison Walvis-Rio, Walvis Ridge and Rio Grande Rise, SMAR, southern Mid-Atlantic Ridge, other abbreviations as in Fig 1 Data from refs 3, 10, 19, 25, 26, 31-34 and A P le R (unpublished data)

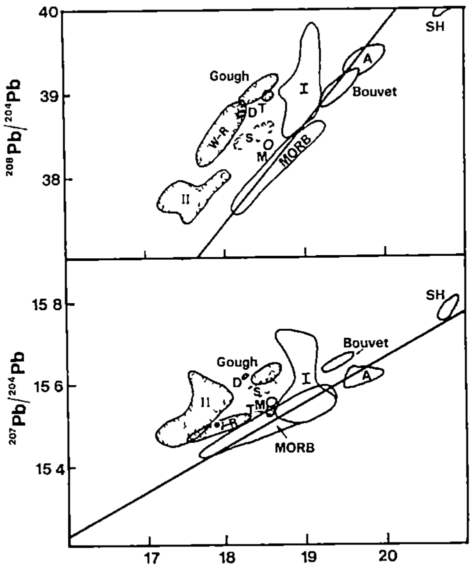


Fig. 3 Pb isotope variation in South Atlantic basalts, group I and group II kimberlites and MORB Initial isotopic ratios of the kimberlites have been used for plotting NHRL, Northern Hemisphere reference line[3], D, Discovery, T, Tristan, SH, St Helena, other abbreviations as in Figs 1 and 2 Data from refs 1, 2, 19, 21 and S R Hart (unpublished)

Zr/Nb, Ba/Nb and La/Nb are highly incompatible trace element ratios that are not normally fractionated during partial melting processes However, the very low degrees of partial melting possibly associated with alkaline, and particularly kimberlite, magma generation require that some caution be exercised when interpreting variations in such inter-element ratios Nevertheless, the good correlations exhibited between these ratios and Sr, Nd and Pb isotopic ratios suggest that the variations shown in Figs 1 and 2 reflect source region composition and are not primarily due to variations in degree of partial melting or residual source mineralogy

Although the trends depicted in Fig 1 correspond broadly to mixing between a depleted MORB-type reservoir and highly enriched reservoirs (reflected by group I and group II kimberlites as possible end-members), Figs 2 and 3 negate any simple mixing relation between the Dupal hotspot-group II kimberlite trend and a depleted MORB-type source Even within the Dupal hotspot-group II kimberlite trend, relations are complex and there is, for example, no simple correlation between ^{87}Sr/^{86}Sr and ^{206}Pb/^{204}Pb along the Pb-Pb arrays[4]

Before evaluating possible implications of the geochemical correlations shown in Figs 1-3, it is necessary to consider current models for the origin of hotspot volcanism Allègre and Turcotte[11] recently summarized current ideas on hotspot source regions and have proposed an integrated model which recognizes two 'end-member' hotspot types The first derives from upwelling of primordial or near-primordial mantle across the asthenosphere-mesosphere boundary, the second from upwelling from a 'mesosphere boundary layer' of recycled (subducted) oceanic lithosphere including pelagic sediments[12,13] and/or delaminated sub-continental lithosphere[14] Delaminated lithosphere may alternatively occur as 'rafts' within the asthenosphere rather than as a boundary layer[15] Evidence for a primordial source component derives mainly from the high ^{3}He/^{4}He isotopic ratios associated with hotspots such as Hawaii and Iceland[16], and evidence for a recycled component derives from the higher ^{208}Pb/^{204}Pb and ^{207}Pb/^{204}Pb ratios and low ^{3}He/^{4}He ratios of hotspots such as Gough and Tristan da Cunha[11,16] The sub-parallel Pb isotope array of Dupal hotspots also requires that the recycled material composing this mesospheric boundary layer is ancient (>1 Gyr)[4,11]

In the South Atlantic, Bouvet (with a high ^{3}He/^{4}He ratio[17]) can be regarded as an example of a hotspot with a significant primordial component in its source region[18], whereas Gough, Tristan da Cunha, Discovery and Shona appear to be hotspots with a major recycled component in their source regions[11,16,18,19] This interpretation is also consistent with the trace element variations shown in Fig 2, where Gough, Tristan and Discovery define a trend extending away from compositions typical for normal ocean island basalt (containing a major primordial component), such as Bouvet, Marion and Ascension, towards high Ba/Nb and La/Nb ratios characteristic of recycled oceanic lithosphere plus sediment or of sub-continental lithosphere[20,21]

Trace element and isotopic variations in South Atlantic hotspots therefore suggest that the Dupal anomaly reflects an area where ancient recycled material exists within the asthenosphere Upwelling from this material has resulted in hotspot volcanism with an apparent lithospheric signature, whereas upwelling outside this region has led to hotspot volcanism with a more normal ocean island basalt signature

It is remarkable that in all three figures (Figs 1-3), group I and group II kimberlites plot either at the end of, or are superimposed on, the non-Dupal hotspot and Dupal hotspot trends respectively Group I and group II kimberlites therefore appear to be derived from source regions similar to the enriched components invoked respectively for non-Dupal and Dupal hotspots In group II kimberlites and Dupal hotspots this component has distinctively low ^{206}Pb/^{204}Pb and high ^{208}Pb/^{204}Pb, Ba/Nb and La/Nb ratios In group I kimberlites and non-Dupal hotspots this component has high ^{206}Pb/^{204}Pb and low Ba/Nb and La/Nb ratios

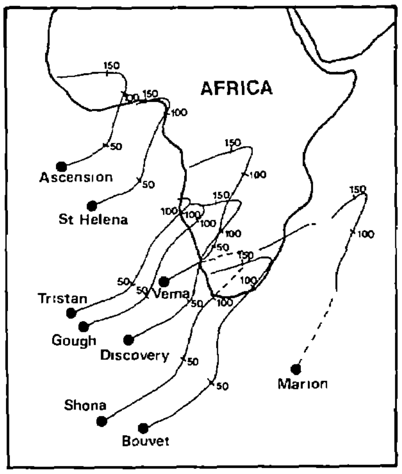


Fig 4 Diagram modified from ref 30 showing South Atlantic hotspots and their predicted palaeo-tracks The Shona hotspot and associated track is from ref 9 Numbers on tracks show ages in Myr

Two separate source regions therefore occur within the asthenosphere beneath the South Atlantic capable of generating alkaline volcanism with essentially the same distinct geochemical signatures as shown by group I and group II kimberlites We may extend this observation to recent plate-tectonic models advocating a correlation between the palaeo-position of hotspots and kimberlite magmatism[22-24]

The African plate contains many hotspots, several of which appear to record the entire history of the opening of the South Atlantic (Fig 4) On the basis of plate reconstructions, Cretaceous and Jurassic kimberlites in southern Africa have been explicitly related to the Bouvet and 'Meteor' (recently relocated and renamed Shona[9]) hotspots[23,24] Geochemical evidence presented here is consistent with such a model in that upwelling from these (or related) hotspots, when located beneath southern Africa (Fig 4), could lead to the deep-seated melting and subsequent eruption of kimberlite magmas with inherently different geochemical signatures, reflecting those of their respective source regions Kimberlites derived from upwelling associated with the Bouvet (or similar) hotspot would have group I geochemical characteristics, whereas those derived from upwelling associated with the Shona (or similar) hotspot would have group II characteristics In this regard, note that kimberlites in Ghana, Guinea and Sierra Leone have group I characteristics, which is consistent with their location on the palaeo-track of the Ascension hotspot (J J Gurney, personal communication, Fig 4), which has essentially the same trace element and isotopic signature as Bouvet (Figs 1-3)

Recent plate reconstructions[9,30] are consistent with a relationship between the Shona hotspot and group II (114-150 Myr) kimberlites but are not readily compatible with a simple relationship between Cretaceous group I kimberlites and the Bouvet hotspot (Fig 4) Group I kimberlites are too young to be directly associated with Morgan's[30] Bouvet palaeo-track and it is possible that in a more complex model, refinements to the plate-tectonic reconstructions, or the presence of an unrecognized non-Dupal hotspot located to the north-east of Bouvet will be required to resolve this problem

The model proposed here places the source of both group I and group II kimberlites within the asthenosphere, with group II kimberlites deriving from recycled lithospheric material This avoids having to derive spatially juxtaposed kimberlite magmas of similar age from two tectonically different source regions Details relating specific hotspot tracks to specific kimberlite pipes still need to be resolved However, the results presented here document the presence of a source region, located within the oceanic asthenosphere, capable of generating alkaline volcanism with essentially the same lithospheric signature as shown by group II kimberlites, so lending credence to models whereby kimberlite magmatism in general is related to the palaeo-positions of postulated mantle hotspots

I thank the University of Cape Town and the Foundation for Research Development for financial support, C B Smith for discussions, J J Gurney, A M Reid and S R Hart for critical comments, S R Hart for allowing me to use his unpublished data, and the British Museum for a sample of Discovery seamount

Received 3 June, accepted 29 July 1986

1 Smith, C B *Nature* **304**, 51-54 (1983)
2 Fraser, K J, Hawkesworth, C J, Erlank, A J, Mitchell, R H & Scott-Smith, B H *Earth planet Sci Lett* **76**, 57-70 (1985)
3 Smith, C B, Gurney, J J, Skinner, E M W, Clement, C R & Ebrahim, N *Trans geol Soc S Afr* **88** (in the press)
4 Hart, S R *Nature* **309**, 753-757 (1984)
5 Dawson, J B in *Ultramafic and Related Rocks* (ed Wyllie, P J), 241-251 (Wiley, New York, 1967)
6 Smith, C B, Allsopp, H L, Kramers, J D, Hutchinson, G & Rodic, J D *Trans geol Soc S Afr* **88** (in the press)
7 Kramers, J D *Earth planet Sci Lett* **34**, 419-431 (1977)
8 le Roex, A P, Dick, H J B, Reid, A M, Gulen, L & Erlank, A J *Geochim cosmochim Acta* (in the press)
9 Hartnady, C J H & le Roex, A P *Earth planet Sci Lett* **75**, 245-257 (1985)
10 Kempe, D R C & Schilling, J G *Contr Miner Petrol* **44**, 101-115 (1974)
11 Allegre, C J & Turcotte, D L *Geophys Res Lett* **12**, 207-210 (1985)
12 Hofmann, A W & White, W M *Earth planet Sci Lett* **57**, 421-436 (1982)
13 Ringwood, A E *J Geol* **90**, 611-643 (1982)
14 McKenzie, D & O'Nions, R K *Nature* **301**, 229-231 (1983)
15 Hawkesworth, C J, Mantovani, M, Taylor, P N & Palacz, Z *Nature* **322**, 356-359 (1986)
16 Kurz, M, Jenkins, W J & Hart, S R *Nature* **297**, 43-47 (1982)
17 Kurz, M thesis, Woods Hole Oceanogr Inst (1982)
18 le Roex, A P in *Mantle Metasomatism* (eds Menzies, M A & Hawkesworth, C J) (Academic, London, in the press)
19 le Roex, A P *J Petrol* **26**, 149-186 (1985)
20 Hole, M J, Saunders, A D, Marriner, G F & Tarney, J *J geol Soc Lond* **141**, 453-472 (1984)
21 Hawksworth, C J, van Calsteren, P, Rogers, N W & Menzies, M A in *Mantle Metasomatism* (eds Menzies, M A & Hawkesworth, C J) (Academic, London, in the press)
22 Duncan, R A, Hargraves, R B & Brey, G P *Geol Mag* **115**, 317-327 (1978)
23 Crough, S T, Morgan, J W & Hargraves, R B *Earth planet Sci Lett* **50**, 260-274 (1980)
24 Duncan, R A *Tectonophysics* **74**, 29-42 (1981)
25 le Roex, A P *et al* *J Petrol* **24**, 267-318 (1983)
26 le Roex, A P & Erlank, A J *J Volc geotherm Res* **13**, 309-338 (1982)
27 O'Nions, R K, Hamilton, P J & Evensen, N M *Earth planet Sci Lett* **34**, 13-22 (1977)
28 Sun, S S *Phil Trans R Soc* **A297**, 409-445 (1980)
29 Erlank, A J *et al* in *Mantle Metasomatism* (eds Menzies, M A & Hawkesworth, C J) (Academic, London, in the press)
30 Morgan, J W *Tectonophysics* **94**, 123-139 (1983)
31 Kable, E J D thesis, Univ Cape Town (1972)
32 Thompson, G, Humphris, S & Schilling, J G *Init Rep DSDP* **72**, 457-466 (1983)
33 Humphris, S & Thompson, G *Earth planet Sci Lett* **66**, 223-242 (1983)
34 Harris, C, Bell, J D & Atkins, F B *Earth planet Sci Lett* **60**, 79-85 (1982)

Archaean flow-top alteration zones formed initially in a low-temperature sulphate-rich environment

Donald R. Lowe & Gary R. Byerly

Department of Geology, Louisiana State University,
Baton Rouge, Louisiana 70803, USA

The 3,500–3,400 Myr Onverwacht Group in the southern part of the Barberton Greenstone Belt, South Africa, includes up to 12 km of komatiitic and tholeiitic lavas containing minor sedimentary and felsic volcanic units[1]. Komatiitic lavas showing spinifex textures, chrome spinels and high Ni and Cr contents are common throughout the sequence[2]. The rocks have been subject to polyphase Archaean deformation and low-grade metasomatism[2-4]. At several positions within the Onverwacht Group (Fig. 1), the tops of komatiitic lavas underlying chert beds are marked regionally by 1–50-m thick zones of silicification, carbonatization and K_2O enrichment (Fig. 2). Because these zones have been variously interpreted to represent the felsic tops of mafic-to-felsic volcanic cycles, weathering horizons and shear zones, their origin is germane to any interpretation of the stratigraphy, structure and evolution of the Barberton belt. Here we present evidence that they formed in

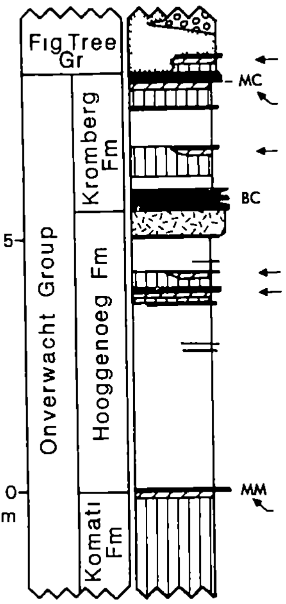


Fig 1 Generalized stratigraphic column of the Onverwacht and lower Fig Tree Groups in the southern part of the Barberton Greenstone Belt showing the distribution of komatiitic flow top alteration zones (arrows) Symbols vertical rule, komatiitic volcanic units, diagonal rule, flow-top alteration zones, unpatterned, tholeiitic units, random line segments, felsic units, black, cherts, stipple and circles, other sedimentary units, MM, Middle Marker BC, Buck Ridge Chert and MC, Msauli Chert

large part by low-temperature, near-surface alteration of komatiitic lavas during intervals of volcanic quiescence. They do not mark fault zones and cannot be used to subdivide the greenstone belt sequence into a series of fault-bounded structural slices.

The altered rock (Fig 3*a*) consists of green, silicified komatiite cut by several generations of veins but, most distinctively, by anastomosing, crudely stratiform veins of silica and carbonate[2] Komatiitic remanents have been altered to microcrystalline aggregates that can include chlorite, sericite, chrome-rich sericite, talc, carbonate and chert, but many show well-preserved silica pseudomorphs of igneous spinifex and cumulate minerals[2] Of primary igneous minerals, only chrome-rich spinels, which have compositions identical to those in less altered komatiites, have survived relatively intact Shearing effects are developed where the zones are cut by later faults, but over wide areas, primary igneous textures are finely preserved and there is no evidence of shearing, cataclasis or mineral foliation Overall, these are zones of extreme SiO_2 and K_2O enrichment and MgO and FeO depletion[5,6] (Table 1)

Crudely stratiform veins, ranging from <1 mm to >5 cm thick, compose as much as 75% of the rock (Fig 3*a*) They consist of microcrystalline quartz (chert) and ankeritic carbonate that commonly display a fine cross-vein fibrous structure (Fig 3*a*) Most fibres are straight, although not necessarily perpendicular to vein walls, but some are sigmoidal The veins contain or are separated by chunks of wall rock, and many are bounded above and below or truncated laterally by non-fibrous cavity-filling chert and druzy quartz (Fig 3*a*) These veins apparently developed by crack-sealing in response to vertical extension[7,8]

In many areas, the altered rock contains quartz nodules (Fig 3*b*) These include small quartzine (length-slow chalcedony) nodules, generally <2 mm in diameter, and large nodules, up to 5 cm in diameter, consisting of curved, radiating, plumose aggregates of length-slow pseudofibrous or flamboyant quartz

The textures of the vein-filling chert and carbonate indicate that they are pseudomorphs of earlier fibrous constituents The following features suggest that one of these constituents was gypsum (1) Within the Onverwacht Group, primary grains of coarse quartz, such as quartz phenocrysts in felsic volcanic rocks and volcaniclastic sandstones and cavity-filling druzy quartz, are preserved without alteration The microcrystalline character of the fibre-replacing quartz suggests replacement of a more poorly ordered silica type, such as chalcedony, or of a non-quartzose mineral, such as carbonate or gypsum (2) The widespread dissolution of the fibrous veins to form cavities later filled by non-fibrous silica suggests that the original vein material was highly soluble, unlike quartz and ankeritic carbonate (3) As noted by Ramsay and Huber[7], 'the length of the fibre [in extensional veins] generally has no special orientation with respect to the crystallographic directions of the crystal making up the fibre ' Virtually all fibre-replacing quartz in the alteration zones, however, shows length-slow aggregate polarization suggesting that it is a pseudomorph of a fibrous length-slow precursor, probably chalcedony Microcrystalline quartz in the adjacent wall rock is randomly oriented The associated quartz nodules also consist of length-slow fibrous and pseudofibrous quartz Whereas most chalcedony is length fast, chalcedony replacing sulphate and evaporites is characteristically length slow[9] Quartzine spherules and flamboyant quartz are also common silica replacement products of anhydrite nodules[10,11] We suggest that stratiform veins in these alteration zones were formed as satinspar or satinspar-calcite crack-sealing veins, similar to fibrous gypsum and carbonate veins widely developed at shallow levels in Phanerozoic sequences[12], and that the quartz nodules represent silicified anhydrite nodules

These distinctive alteration zones are best developed at the tops of komatiitic flow or pyroclastic sequences that are overlain directly by chert units >1 m thick Although zones of silicification and K_2O enrichment occur beneath cherts at the tops of tholeiitic and felsic sections, networks of stratiform veins and silica nodules are present only where the host rock is komatiite Primary host-rock lithology, eruptive style or eruptive setting therefore played a critical role in their formation

The upper contacts of the alteration zones are generally flat but locally show up to 6 m of erosional relief Overlying silicified sediments commonly rest with depositional contact on altered komatiite and contain komatiitic clasts and detrital chrome-rich spinels Beneath depositional contacts, the uppermost 10–50 cm

Table 1 X-ray fluorescence analyses of altered and fresh komatiitic flow rocks, Barberton Greenstone Belt

	1	2	3	4
SiO_2	85 61	52 38	82 52	52 35
Al_2O_3	8 30	7 16	9 44	10 64
FeO^T	1 64	12 40	1 95	10 57
MgO	1 07	16 33	1 79	12 79
CaO	0 18	9 37	0 46	11 48
Na_2O	0 05	1 11	0 08	1 14
K_2O	2 45	0 08	2 98	0 05
TiO_2	0 53	0 69	0 76	0 59
P_2O_5	0 12	0 08	—	0 05
MnO	0 01	0 23	0 03	0 20
Cr	2,399	1,486	4,387	2,007
Ni	188	289	849	348
Zr	41	48	64	38
Y	21	16	12	14

All Fe expressed as FeO, dry weight normalized to 100%

1 Altered coarse-spinifex komatiite from uppermost alteration zone shown in Fig 1 Analysis by J Veizer and R Hartree, Department of Geology, University of Ottawa

2 Fresh coarse-spinifex komatiite from Fig Tree Group Analysis by G Byerly, Department of Geochemistry, University of Cape Town

3 Altered komatiitic basalt from near middle of Kromberg Formation, Komati River Analysis by J Veizer and R Hartree

4 Fresh basaltic komatiite below sample 3 Analysis by G Byerly

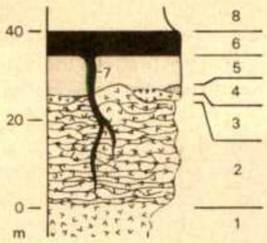


Fig. 2 Idealized sequence of lithologic units associated with alteration zones at the tops of komatiitic flow units. The areal distribution of alteration zones has been outlined previously (Fig. 3b in ref. 2). (1) serpentinized komatiite, (2) alteration zone with crudely stratiform veins, (3) breccia at top of alteration zone, (4) lenticular units of coarse-grained, poorly sorted komatiitic sandstone, (5) silicified volcaniclastic and pyroclastic deposits, (6) black and banded chert, (7) chert dikes, and (8) overlying volcanic or sedimentary units.

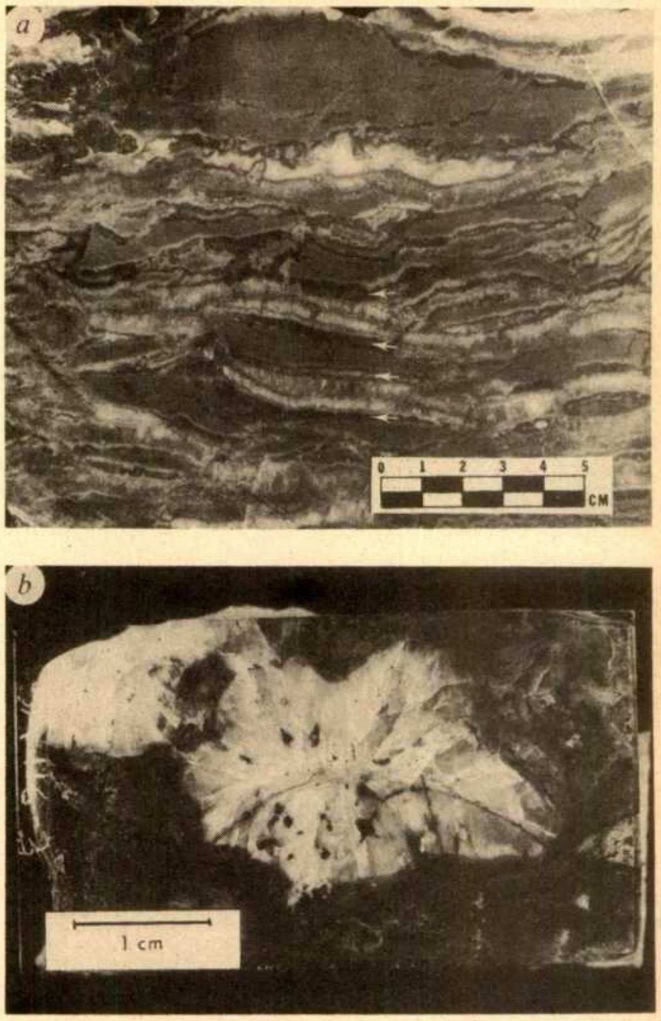


Fig. 3 *a*, Altered komatiitic flow rock showing development of crudely stratiform, fibrous veins. Medium-gray granular masses are altered komatiite remanents. Parallel layers of cavity-filling, non-fibrous chert (dark) and druzy quartz (white) adjacent (arrows) to many veins suggest at least two generations of cracking and vein development. *b*, nodules of plumose flamboyant length-slow quartz in silicified komatiite from alteration zone in the lower part of the Fig Tree Group.

of komatiite are usually brecciated (unit 3, Fig. 2). The breccias consist of angular clasts of unveined komatiite separated by dilatational chert-filled fractures that commonly contain volcaniclastic debris and carbonaceous matter similar to overlying sedimentary units. These breccias are unlike volcanic flow-top breccias and appear to have formed by *in-situ* dilatational fracturing.

Capping cherts (unit 5, Fig. 2) include silicified carbonaceous and volcaniclastic layers deposited mainly under low-energy, shallow-water conditions[13,14]. Locally, lenses of coarse-grained, poorly sorted, cross-stratified komatiitic debris (unit 4) or stromatolites in growth position rest directly on the komatiites or komatiitic breccias, reflecting more energetic shallow-water to emergent conditions[15]. The absence of chunks of stromatolite and sedimentary rock in the breccias indicates that brecciation occurred before or during deposition of the immeidately overlying sediments. Many sedimentary layers contain pseudomorphs after crystals of early diagenetic gypsum[13].

The upper parts of these sedimentary layers generally include units of black chert representing silicified carbonaceous sediments (unit 6, Fig. 2). In several areas, chert dikes (unit 7), formed by the intrastratal flowing of fluid carbonaceous oozes, extend downward from these units, through underlying sedimentary layers, and into the alteration zones. No dikes have been observed extending upward from black chert layers. These dikes truncate veins in the alteration zones, indicating that vein formation preceded diking. Both dikes and veins apparently formed before deposition of more than a few tens of metres of suprajacent sediment.

These alteration zones have been variously interpreted as altered silicic tuffs[1,16], cataclastic units marking faults[3,8], zones of relatively high-temperature hydrothermal alteration[5,6], and weathering zones[2]. Preserved igneous textures and magmatic chrome spinels indicate that the original rocks were komatiitic, not felsic volcanics[2]. Indications that the formation of these zones involved little or no shearing are: the presence of abundant, unstrained spinifex (1), the common absence of mineral foliation and cataclasis (2), their regionally stratiform non-discordant geometry over outcrop distances of up to 25 km (3), their widespread depositional contact with suprajacent sedimentary units (4) and the local presence of earlier-formed vertical veins cut but not offset by the stratiform veins (5) (Fig. 6*b* of ref. 2).

Our results suggest that these alteration zones formed through the following sequence of partly overlapping events. (1) The rapid eruption of thick sequences of komatiitic lavas in overall shallow-water settings resulted in a widespread build-up of the volcanic surfaces into the subaerial environment. (2) Post-eruption subaerial exposure of the flow tops was accompanied by erosion and the local accumulation of komatiitic alluvial sand and gravel (unit 4, Fig. 2). This may also have been a time of chemical weathering of the flow tops. (3) Regional subsidence in the absence of volcanism led eventually to submergence. Passage of the flow tops through marginal evaporitic environments was associated with planation, formation of anhydrite nodules, brecciation, and local growth of stromatolites. (4) Shallow-water deposition of volcaniclastic and carbonaceous sediments resulted in infiltration of the breccias by loose debris and the accumulation of sediment caps on the komatitic flows. (5) Silicification of fine-grained volcaniclastic units began almost concurrently with deposition[13,17], forming nearly impermeable chert caps above the alteration zones. Silicification of the upper parts of the lavas may also have begun at this time. Fluids moving through the volcanic rocks were trapped beneath these seals resulting in hydraulic overpressuring and formation of crack-seal veins. These were low-temperature fluids, probably less than 60 °C, as suggested by the precipitation of gypsum[18]. They may have been dense sulphate-rich brines originating in marginal evaporitic environments. (6) The formation of chert dikes reflects fracturing of the chert caps above the alteration zones and may mark the release of hydraulic overpressures and the end of vein formation. Later changes in pore-water chemistry, perhaps related to burial or the upward movement of deeper-level, higher-temperature fluids, resulted in gypsum replacement and later stages of silicification, carbonatization and K_2O enrichment. Where the chert caps remained intact, the formation of additional crack-seal veins may have accompanied these metasomatic events.

The results of this study indicate that these distinctive Archaean alteration zones formed initially through a complex sequence of surface and near-surface low-temperature processes. Later metasomatism and metamorphism also played a major role in their final character. Studies of this and other flow-top alteration zones may provide important clues to Archaean water and atmospheric chemistry, surface and subsurface hydrologic systems, and the relationships among volcanism, subsidence and sedimentation in greenstone belts.

This research was supported by NSF grants EAR 79-19908, EAR 82-06015 and EAR 84-06420 to D.R.L. and EAR 79-19909 to G.R.B. and D.R.L. We thank the Department of Geology, University of the Witwatersrand, and the Anglo American Corporation of South Africa, for field support.

Received 27 March; accepted 29 July 1986.

1. Viljoen, M. J. & Vilgoen, R. P. *Spec. Publs geol. Soc. S. Afr.* **2,** 113–152 (1969).
2. Lowe, D. R., Byerly, G. R., Ransom, B. L. & Nocita, B. W. *Precambr. Res.* **27,** 165–186 (1985).
3. deWit, M. L. *J. struct. Geol.* **4,** 117–136 (1982).
4. deWit, M. J., Hart, R., Martin, A. & Abbott, P. *Econ. Geol.* **77,** 1783–1801 (1982).
5. Pearton, T. N. in *Komatiites* (eds Arndt, N. E. & Nisbet, E. G.) 459–476 (Allen & Unwin, London, 1982).
6. Hanor, J. S. & Duchac, K. C. *Absts Prog. geol. Soc. Am.* **17,** 603.
7. Ramsay, J. G. & Huber, M. I. *The Techniques of Modern Structural Geology Vol. 1: Strain Analysis,* 235–263 (Academic, London, 1983).
8. deWit, M. J. in *Abstr. Workshop tecton. Evol. Greenstone Belts* **25,** (Lunar & Planetary Institute, Houston, 16–18 Jan. 1986).
9. Folk, R. L. & Pittman, J. S. *J. sedim. Petrol.* **41,** 1045–1058 (1971).
10. Chowns, T. M. & Elkins, J. E. *J. sedim. Petrol.* **44,** 885–903 (1974).
11. Tucker, M. E. *Sedim. Geol.* **16,** 193–204 (1976).
12. Machel, H. *Sedimentology* **32,** 443–454 (1985).
13. Lowe, D. R. & Knauth, L. P. *J. Geol.* **85,** 699–723 (1977).
14. Lanier, W. P. & Lowe, D. R. *Precambr. Res.* **18,** 237–260 (1982).
15. Byerly, G. R., Lowe, D. R. & Walsh, M. M. *Nature* **319,** 489–491 (1986).
16. Anhaeusser, C. R. *Phil. Trans. R. Soc.* **A273,** 359–388 (1973).
17. Paris, I., Stanistreet, I. G. & Hughes, M. J. *J. Geol.* **93,** 111-129 (1985).
18. Hardie, L. A. *Am. Miner.* **52,** 171–200 (1967).

Body size, ecological dominance and Cope's rule

James H. Brown & Brian A. Maurer*

Department of Ecology and Evolutionary Biology,
University of Arizona, Tucson, Arizona 85721, USA

We present data and analyses demonstrating that large species utilize a disproportionately large share of the resources within local ecosystems. Even though small species tend to have higher local population densities, these are not sufficient to compensate for their lower rates of energy use per individual. The relationship is very general; holding for example for birds, mammals, fish and plants. We suggest that several ecological advantages enable larger species and larger individuals within species to monopolize resources, and that the resulting selection pressures are responsible for the evolutionary trend towards increasing body size seen in many phyletic lineages. Our results contradict important studies[1–4] that have concluded that species of small body size use at least as large a proportion of the resources within ecosystems as their larger relatives.

A goal of community ecology is to understand the patterns and processes that determine the numbers and variety of coexisting species. Are there general rules that characterize the contributions of the different species to the structure and dynamics of ecosystems? One measure of the ecological importance of species is their relative utilization of energy and transfer of materials[5,6]. Within groups of taxonomically-related, ecologically similar organisms, two variables (individual body size and population density) are of primary importance in determining how energy and other limiting resources are allocated. Large individuals demand more from environmental resource pools than small ones, but also tend to occur at lower population densities. The question, then, is to what extent the greater abundance of small species compensates for their lower resource requirements per individual.

Although direct measurements of energy flow per unit area are probably not available for all species in any community, these rates can be estimated quite accurately from data on population density (D), individual body mass (M) in animals and photosynthetic cover (C) in plants. In animals the average daily energy use of a free-living individual (U) scales allometrically as approximately $M^{0.67}$ (refs 2, 7–10), so total energy use per species (E) is given as $E = DU = DM^{0.67}$. In plants, variation in photosynthetic rate per unit leaf surface among species is insignificant compared to the total leaf surface[11,12], so the energy use of each species is directly proportional to its vegetative cover (C).

We analysed data on D and M for local groups of birds, mammals and fish, and on C for communities of plants to determine how rates of energy use per species (E) varies with body size. Using the results of 39 censuses of breeding land birds and data on species body mass[13], we tried three different methods[14] to fit a linear relationship for log D with log M (Table 1). Depending on the method, only 3 to 10 of the 39 slopes were ≤ -0.67. Pooling the data gave a slope of approximately -0.30, not greatly different from the value of -0.19 reported by Peters[2]. If D scales as $M^{-0.30}$ and U scales as $M^{0.67}$, then the rate of energy intake per species ($E = DU$) increases with body mass, scaling as $\sim M^{0.37}$. Because the three methods of estimating slopes give comparable results, in the following analyses we present only the results of the regression method applied to log-transformed data.

For the nine species of granivorous rodents (mammals) in a Chihuahuan Desert ecosystem[15], regression of log D on log M gave a significant positive relationship ($r = 0.56$, slope $[b] = 1.17$, $n = 9$, $P = 0.05$). Consequently, in these mammals the rate of energy flow per species ($E = DU$) scales positively with body mass as $\sim M^{1.17} \times M^{0.67} = M^{1.84}$.

We analysed three independent data sets[16–18] on population densities and body masses of marine fish (Table 2). The slopes of regressions of log D on log M averaged near zero, and only one was more negative than -0.1. Thus, if D scales as $\sim M^{0.0}$ and U scales as $M^{0.67}$, then the average rate of energy use per species ($E = DU$) must vary as $\sim M^{0.67}$.

We obtained complete census data for perennial plants in five different habitats in two desert regions[19,20]. Total cover (C) of each species was used to estimate its contribution to energy flow, and cover divided by population density (D) was taken as the effective size of an individual plant. When we regressed log C against log (C/D), all the slopes were positive, with values ranging from 0.53 to 1.59 and averaging ~ 1.0 (Table 3). Although the usefulness of estimating individual mass in perennial plants containing nonliving woody tissue is questionable, M should vary with individual surface area as $\sim (C/D)^{1.5}$ and rate of energy use per species (E) should scale as $> M^{1.0}$.

Our analyses show that in four different taxonomic groups (birds, mammals, fish and plants) species comprising large individuals account for most of the energy flow and resource use within local ecosystems. These results contradict those of two influential recent studies. Damuth[1] claimed that population densities of mammals scale as $M^{-0.75}$, so that rates of energy use per species are essentially independent of body size (E scales as $\sim M^{0.0}$). Peters[2–4] compiled data which suggest that population density (D) often scales as $M^{-1.0}$ so that energy use per

---

* Present address: Department of Zoology, Brigham Young University, Provo, Utah 84602, USA.

Table 1 Slopes of log population density and log body mass[13] for land birds in different habitats

Habitat	log data regression	log normal data regression	log normal data RMA	Correlation coefficient	No. of species
Savanna	−0.53	−0.64	−1.11	−0.40	17
Grassland	0.19	−0.18	1.36	0.45	19
Deciduous forest	1.08	2.47	7.29	0.48	15
Managed forest	−0.05	−0.25	−1.40	−0.04	35
Douglas fir	−0.90	−1.03	−1.93	−0.57	21
Clearcut forest	−0.31	−0.99	−2.86	−0.26	15
Ponderosa pine	−0.18	−0.18	−0.32	−0.26	21
Logged ponderosa	−0.44	−0.57	−0.60	−0.47	22
	−0.31	−0.33	−0.44	−0.39	23
	−0.04	−0.07	−0.10	−0.06	19
Clearcut forest	−0.72	−0.18	−0.12	−0.50	6
Spruce	−0.41	−0.20	−0.20	−0.25	16
	−0.51	−0.17	−0.09	−0.34	21
Spruce/jackpine	−0.58	−0.31	−0.37	−0.32	22
Spruce forest edge	−0.60	−0.32	−0.21	−0.35	12
Spruce/jackpine edge	−0.59	−0.35	−0.44	−0.31	21
Spruce riparian	−0.50	−0.46	−0.40	−0.24	18
Beech/maple	−0.37	−1.34	−1.62	−0.41	29
	−0.37	−1.51	−1.23	−0.51	24
Urban woodlot	0.39	0.73	5.15	0.36	20
	0.17	1.73	−0.28	0.20	17
Pine/oak seral	−0.18	−0.15	−0.51	−0.09	23
Pine/oak mature	−0.67	−0.66	−2.32	−0.31	23
Cove forest	−0.17	−0.14	−0.35	−0.11	31
Hemlock/deciduous	−0.63	−0.33	−0.28	−0.45	26
Chestnut oak	−0.36	−0.13	−0.13	−0.23	32
Red oak	−0.31	−0.10	−0.09	−0.29	25
Pine heathland	−0.57	−0.13	−0.09	−0.39	14
Gray beech	−0.08	−0.17	−0.20	−0.06	13
Spruce/fir early	−1.05	−5.68	−5.14	−0.35	11
Spruce/fir middle	−0.34	−4.37	−4.34	−0.25	10
Spruce/fir late	0.10	−0.24	0.34	0.07	14
Spruce/fir climax	−0.37	−0.60	−0.54	−0.23	15
Grassland	−0.01	−0.13	−0.26	−0.01	13
Brushland	−0.49	−0.27	−0.24	−0.40	14
Savanna	0.02	−0.11	0.56	0.01	14
Shrubland	0.03	−0.31	1.22	0.02	15
Aspen	−0.11	−0.02	−0.01	−0.13	28
Savanna	−0.17	−0.44	−1.34	−0.11	12
Pooled data	−0.30	−0.12	−0.40	−0.27	197

Three methods of calculation were used: linear regression on log-transformed data, linear regression and reduced major axis (RMA) fitted to log-normally distributed data[14]. Note that only 3–10 slopes were < -0.67.

These 39 censuses were taken from the published literature; the references are available from the authors on request.

Table 2 Regression analyses of log population density on log body mass for fish

Habitat	Slope	Standard error	Intercept	Correlation coefficient	No. of species
Caribbean coral reef[16]	−0.05	0.08	0.76	−0.08	53
Gulf of California tide pool[17]	0.04	0.22	2.26	0.04	25
Newport Bay, California[15]					
bag seine	−0.08	0.36	1.18	−0.05	27
seine/bay	0.32	0.65	1.22	0.12	19
seine/shallows	−0.07	0.98	1.96	−0.03	8
drop net	−0.48	0.35	0.89	−0.37	14
enclosure	0.24	0.46	1.11	0.25	6
otter trawl	0.16	0.12	0.60	0.22	35
gill net	−0.03	0.26	0.88	−0.03	22

Note that none of the slopes are more negative than −0.50.

species should scale negatively with body mass (as $M^{-0.33}$). The results of Damuth and Peters can probably be attributed to a sampling bias inherent in their data collection procedures. Their analyses are based on published values of population densities for species from different, scattered localities. These data probably neglect the numerous rare species of small body size that occur within local ecosystems, because ecologists tend to study conspicuous and abundant species. We have been careful to avoid this and other possible biases by using only complete censuses of local biotas.

We have calculated the slopes of allometric equations to reveal differences between our results and those of Damuth and Peters. However, the low correlation coefficients of our regressions suggest that no allometric equation adequately characterizes the

Table 3 Regression analyses of log cover on log (cover/population density) for perennial plants in five different habitats in two desert regions of southwestern North America

Habitat	Slope	Standard error	Intercept	Correlation coefficient	No. of species
Sonoran Desert[19]					
Goldeneye	0.55	0.18	0.91	0.53	27
Agave	0.53	0.14	0.67	0.56	33
Mojave Desert[20]					
Shadscale	1.32	0.26	3.40	0.76	21
Bladder-sage	1.25	0.17	3.19	0.87	19
Spiny menodora	1.59	0.13	4.36	0.93	23

Note that all the slopes are greater than 0.50.

relationship between population density and body size. Our qualitative result that on average small species use proportionally less energy than large ones can be corroborated using a robust nonparametric technique. We divide the species in a census into those above and below the median body mass for species in that census. Then if population energy use (E) is on average greater for larger than for smaller species, then $\bar{E}_s < \bar{E}_l$, where $\bar{E}$ represents mean energy use of the smaller (s) and larger (l) species. For animals in which individual energy use scales as $M^{0.67}$, we can substitute $\bar{D}\bar{M}^{0.67}$ for $\bar{E}$, so $\bar{D}_s\bar{M}_s^{0.67} < \bar{D}_l\bar{M}_l^{0.67}$. Because the probability that any one census obeys this relationship by chance is 0.5, whenever we have a sufficient number of censuses we can use a binomial test to evaluate the hypothesis that the collection of censuses exhibits the expected pattern. For the bird censuses in Table 1, 35 out of 39 showed the predicted relationship (one-tailed test, $p < 0.0001$).

We have shown that on a per species basis, within trophic levels and taxonomic groups, organisms of large individual size account for most of the energy flow through local ecosystems. This has important ecological and evolutionary implications. There are several ecological advantages of large body size. Large species are usually dominant in interspecific aggression[21], which may result in exclusion of small species from preferred resources; they tend to have more efficient homeostatic mechanisms and greater mobility[2,8,10], so they are able to tolerate a wider range of environmental conditions and to seek out more favourable locations. An increase in size enables an individual to spend less energy per unit biomass on maintenance and to become more efficient at extracting usable energy from low-quality foods[2,22]. Consequently, the same amount of available energy can support a greater biomass of a large species than of a small one. Finally, size may be important in avoiding predators[23]. Our data suggest that these advantages enable large species to dominate the resource allocation in ecosystems, leaving smaller species to divide up the remainder.

The ecological consequences of body size provide a mechanistic explanation for the evolutionary cycle known as Cope's rule[23,24]. Phylogenetic lineages usually begin with relatively small organisms that give rise to larger forms. Ultimately giant forms are produced, but these frequently become extinct, creating new opportunities for another lineage as it in turn evolves organisms of larger size. Presumably the ecological advantage of monopolizing resources provides the selective pressure that promotes evolution of greater size. Individuals of large size are favoured by intraspecific natural selection, because they can dominate resource use and consequently leave more offspring than their smaller relatives. However, this increase in body size is accompanied by a higher probability of extinction, which may be attributed to lower population density, smaller population size, slower population growth and perhaps other factors[23,24]. Thus Cope's rule is the result of two opposing evolutionary processes, one operating at the level of individuals within populations, and the other at the level of species within ecosystems[25].

We thank M. Kurzius and R. Robichaux for technical assistance; W. A. Calder III, J. Damuth, P. Harvey, C. Istock, D. Jablonski, L. Van Valen and an anonomous reviewer for valuable comments; the University of Arizona for computer funds; and the NSF for research support grant BSR 8021535 to J.H.B.

Received 7 October 1985; accepted 26 September 1986.

1. Damuth, J. *Nature* **290,** 699–700 (1981).
2. Peters, R. H. *The Ecological Implications of Body Size* (Cambridge University Press, Cambridge, 1983).
3. Peters, R. H. & Wassenberg, K. *Oecologia* **60,** 89–90 (1983).
4. Peters, R. H. & Raelson, J. V. *Am. Nat.* **124,** 498–517 (1984).
5. Odum, E. P. *Am. Zool.* **8,** 11–18 (1968).
6. Odum, H. T. *Systems Ecology* (Wiley, New York, 1983).
7. Garland, T. *Am. Nat.* **121,** 571–587 (1983).
8. Mace, G. M. & Harvey, P. H. *Am. Nat.* **121,** 120–132 (1983).
9. Walsberg, G. E. in *Avian Biology* vol. 6 (eds Farner, D. S. & King, J. R.) 161–220 (Academic, New York, 1983).
10. Calder, W. A. *Size, Function, and Life History* (Harvard University Press, Cambridge, 1984).
11. Mooney, H. A. *A. Rev. ecol. Syst.* **3,** 315–346 (1972).
12. Schulze, E. D. in *The Encyclopedia of Plant Physiology*, new series, vol. 126, 616–676 (Springer, Berlin, 1982).
13. Dunning, J. B. *Western Bird Banding Assoc. Monogr.* **1,** 1–38 (Eldon, Cave Creek, 1984).
14. Kermack, K. A. & Haldane, J. B. S. *Biometrika* **37,** 30–41 (1950).
15. Brown, J. H. & Munger, J. C. *Ecology* **66,** 1545–1563 (1985).
16. Smith, C. L. & Taylor, J. C. in *Results of the Tektite Program: Ecology of Coral Reef Fishes* vol. 14 (eds Collette, B. B. & Earl, S. A.) 125–170 (Science Bulletin, Natural History Museum of Los Angeles County, 1972).
17. Thomson, D. A. & Lehner, C. E. *J. exp. mar. biol. Ecol.* **22,** 1–28 (1976).
18. Horn, M. H. & Allen, L. G. *California Department of Fish and Game Tech. Rep.* **45,** 1–36 (1981).
19. Turner, R. M. & Brown, D. E. in *Desert Plants* vol. 4 (ed. Brown, D. E.) 181–221 (University of Arizona, Tucson, 1982).
20. Kurzius, M. A. Thesis, Univ. Nevada, Las Vegas, 1981.
21. Morse, D. H. *Behavioral Mechanism in Ecology* (Harvard University Press, Cambridge, 1980).
22. Sibley, R. M. in *Physiological Ecology* (eds Townsend, C. R. & Callow, P.) 109–139 (Blackwell, London, 1981).
23. Stanley, S. M. *Macroevolution* (Freeman, San Francisco, 1979).
24. Bakker, R. T. in *Patterns of Evolution* (ed. Hallam, A.) 431–478 (Elsevier, Amsterdam, 1977).
25. Fowler, C. W. & MacMahon, J. A. *Am. Nat.* **119,** 480–498 (1982).

Mach bands are phase dependent

M. Concetta Morrone, John Ross, David C. Burr & Robyn Owens*

Department of Psychology and * Department of Computer Science, University of Western Australia, Nedlands, Western Australia 6009, Australia

The prevailing explanation of Mach bands, the paradoxical bands of light and dark seen where luminance gradients meet plateaux, is that they are due to lateral inhibition in the visual system[1–4]. This explanation equates Mach bands with distortions in a processed luminance distribution due to selective attenuation of low frequency components. But square waveforms exhibit no Mach bands[2,5,6], although they should also be distorted after processing. Measurements of the contrast required to see Mach bands in trapezoidal waveforms and manipulations of their spectra lead us to conclude that phase relationships between Fourier components are important to the structure we perceive. A model based on the odd and even symmetry of visual receptive fields explains our results.

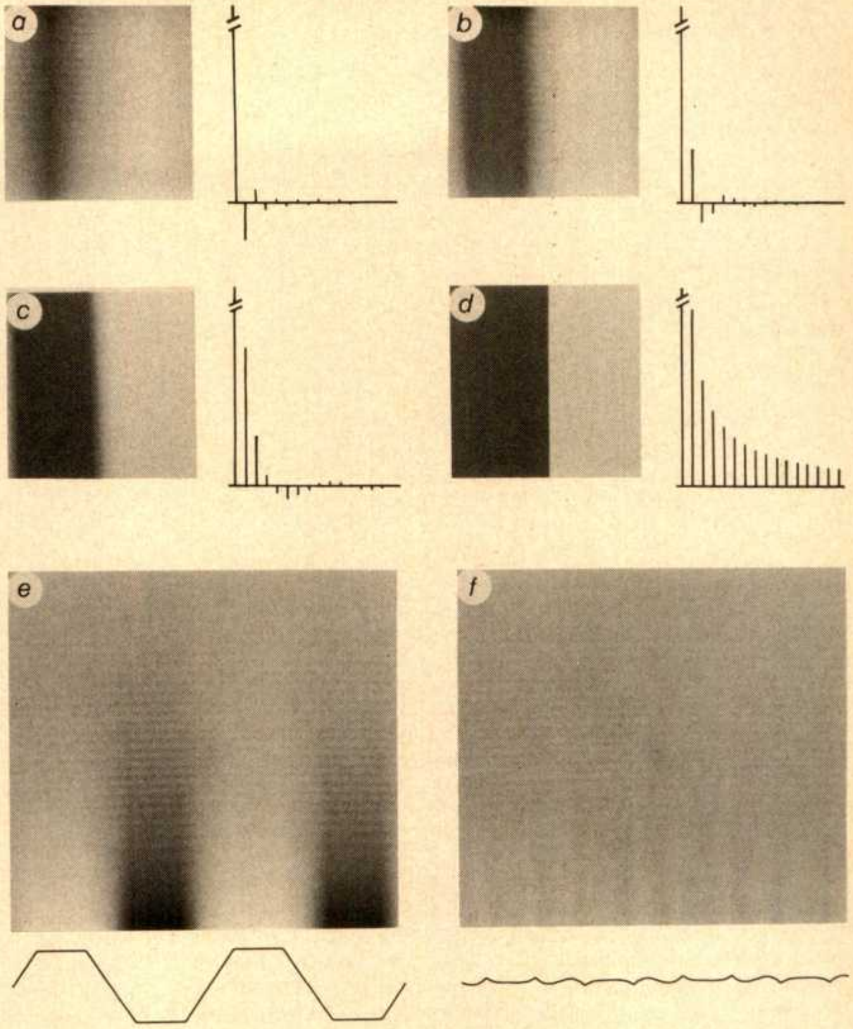


Fig. 1 *a–d*, Examples of four types of trapezoids. The shape of the waveform is determined by the parameter t (the ratio of the ramp width to the period), which is 0.5 in *a* (triangular wave), 0.25 in *b*, 0.125 in *c*, and 0 in *d* (square wave). Mach bands are clearly visible in all except the square wave. The Fourier expansions of the patterns are given by

$$F(x)=\sum_{k=0,\infty}[(4A/t\pi^2)\cdot(\sin(\pi t(2k+1))/(2k+1)^2)]\cdot\sin(2\pi(2k+1)x/T), \qquad (1)$$

where A is the amplitude and T is the period. The spectra are depicted schematically adjacent to each waveform (the first harmonic terms are not to scale). Only the odd terms are non-zero, and they fall into alternatively positively and negatively weighted blocks whose size varies inversely with t. *e, f*, At left a trapezoid waveform ($t=0.25$), decreasing in contrast exponentially upwards. At right the residual waveform, from which the first block of positively weighted harmonics (the first and the third) has been removed, also decreasing in contrast exponentially upwards. For various distances (spatial frequencies), note the point at which the Mach bands disappear from the trapezoid, and also the point at which the residual waveform falls below detection threshold. The two thresholds should be similar. By fixating some distance below the two panels that Mach bands disappear in the periphery as contrast decreases well before the residual ceases to be visible.

Mach originally[7] used rotating discs and drums to study the bright and dark bands that appear where ramps of luminance meet plateaux as, for example, at the borders of the penumbra of a shadow. He later devised optical methods to cast shadows[8]. For the studies reported here we used gratings formed by computer on an oscilloscope which were variable in luminance profile and contrast. Fig. 1 shows four trapezoids, including the two extreme cases of a square wave ($t=0$) and the triangular wave ($t=0.5$), where t is the ratio of the width of each ramp to the period. Classical Mach bands are seen on the trapezoids (Fig. 1*b*, 1*c*); bright bands are visible at the top of each ramp and dark bands at the bottom. Note also that the ramps do not appear to have a uniform brightness gradient and the plateaux do not appear uniformly dark and light. The triangular wave also exhibits sharp bands at the peaks and troughs of the waveform, but the square wave does not.

The Fourier transforms of the trapezoid waveforms (calculated from the point where the waveform crosses zero) are also shown in Fig. 1. The components are arranged in blocks alternatively positively and negatively weighted (that is, differing in phase by 180°). The number of components in each block varies inversely with t (see formula in the legend), being one for the triangular wave, and infinite for the square wave. The observations that Mach bands are not seen on square waves[2,6], which have no out-of-phase components, or on trapezoids with very steep ramps[5], where the out-of-phase components are all of high spatial frequency, led us to suspect that phase may hold the key to Mach bands.

We isolated residual waveforms from trapezoids by removing the first block of positive harmonics (see Fig. 1*f*) and measured the contrast necessary to see them as a function of spatial frequency (inverse of their period, T in equation (1), Fig. 1, legend). The results are shown by the filled circles in Fig. 2. We then measured the minimum contrast at which Mach bands could be seen in the original trapezoidal waveforms (open triangles in Fig. 2). The two curves follow one another closely as spatial frequency is varied. Interestingly, observers spontaneously remarked that the two tasks, setting thresholds for seeing Mach bands and for detecting residual waveforms, seemed to make common demands on them, as if the two tasks were almost identical. Readers can confirm some of the results for themselves by observing the contrast at which the Mach bands of Fig. 1*e* and the pattern of Fig. 1*f* disappear at various viewing distances.

This result suggests that Mach bands are visible only when the out-of-phase Fourier components (together with all higher harmonics) reach their independent threshold. To test this more directly, we attenuated the higher frequency components by applying a low-pass gaussian filter to various trapezoidal waveforms and decreased the space constant of the gaussian filter until the Mach bands disappeared. We then applied the same filter to the residual waveforms. For all the trapezoids we investigated, the low-pass filter which removed the Mach bands attenuated the appropriate residual waveform to a contrast just below its detection threshold.

It is known that phase discrimination deteriorates in the periphery[9,10]. We measured the contrast required to see Mach bands in various trapezoids and the contrast required to see their residuals as a function of retinal eccentricity. As eccentricity increased the contrast required to see Mach bands increased

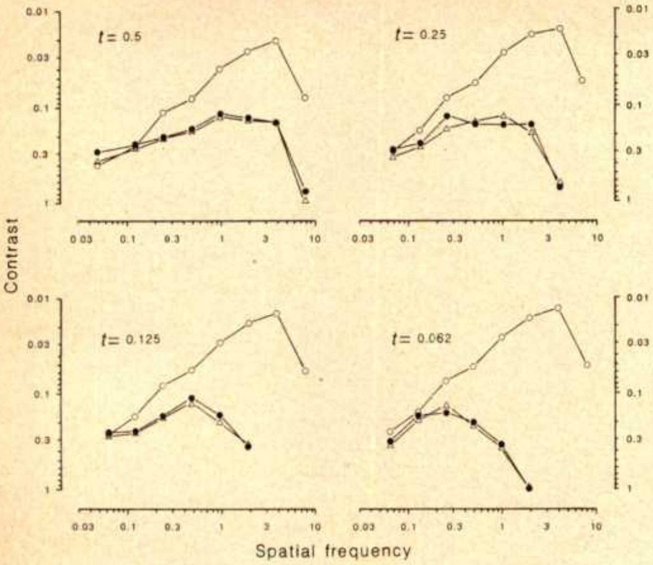


Fig. 2 Contrast required to see Mach bands as a function of spatial frequency (c per deg) is shown by open triangles, for four different t values. Contrast required to see the residual waveform (first block of positive components removed) is shown by the filled circles. The curves follow one another closely. Open circles show contrast required to see the fundamental of the trapezoid, for comparison. All contrasts are expressed as those of trapezoids from which the waveforms derive.

much more rapidly than that required to see the residual. Indeed, Mach bands could not be seen at all at eccentricities greater than three degress (as previously observed by Mach[8]), whereas the residual was still clearly visible. Phase sensitivity is indeed poor in the periphery; when viewed peripherally, trapezoids become unstable and alternate in appearance between square and triangular waves.

As a final confirmation of the importance of phase, we created synthetic waveforms by inverting the sign of all negatively weighted components of the triangular wave ($t = 0.5$) and a trapezoid ($t = 0.25$), to make 'all positive' waveforms, having the same amplitude spectra as the parent trapezoids (Fig. 3). As can be seen, the luminance profiles are now smooth. There are no Mach bands in either case, but sharp edges are apparent, similar to those of a square wave.

The effects of phase on the appearance of distributions of luminance are more radical when there is variation in two dimensions. The distribution of luminance in Fig. 3*c* is pyramidal. In both horizontal and vertical cross-section, the luminance profile is triangular. The Mach bands of one dimensional triangular waveforms reappear in two dimensions as star-like patterns where the two dimensional pattern forms points of high and low luminance. Adjacent is the same pattern with all phases set positive. The structure is now like that of a checkerboard, with white and black squares clearly separated by sharp vertical and horizontal borders.

The results of this study suggest that Mach bands are different from edges, but that both depend on phase relationships. When all negative harmonics (180° out-of-phase) are rendered invisible (by lowering their contrast, increasing their spatial frequency, or smooth digital filtering), Mach bands disappear. When patterns are displayed peripherally (where the visual system is poor at resolving phase[9,10]) the bands also disappear. Finally, if all negative harmonics of a trapezoid are shifted in phase by 180°, Mach bands disappear, but an edge is seen where there was none before.

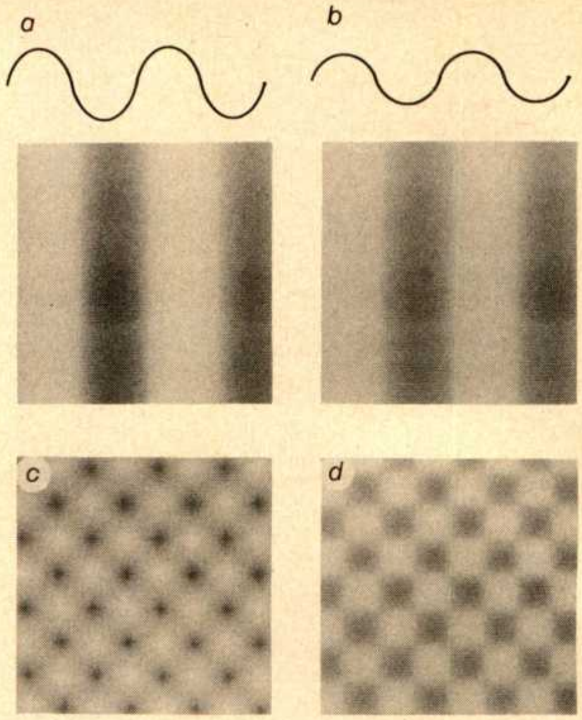


Fig. 3 *a, b*, Synthetic waveforms constructed by inverting the sign of all negatively weighted components of trapezoids of t value 0.5 (*a*) and 0.25 (*b*). The waveforms have no discontinuities, yet we perceive sharp edges, with no Mach bands; *c*, two dimensional pattern formed by multiplying a horizontal by a vertical triangular wave, shows white and black star patterns whch are two-dimensional analogues of Mach bands; *d*, two dimensional pattern derived from *c* by inverting the phase of all negatively weighted harmonics. The star patterns reflecting two-dimensional Mach bands are no longer visible. Instead, a checkerboard-like pattern is seen, with sharp borders between the black and white squares.

As mentioned earlier, all harmonics of a square wave are in zero phase at the point where the waveform crosses zero in a positive direction and 180° where it crosses zero in a negative direction. This is also true for 'all positive' trapezoids in which all out-of-phase components have been shifted in phase by 180°. Elsewhere, the phases of the harmonics advance at different rates, so no other points have strong congruence of phase. Trapezoidal waveforms other than the square wave have blocks of negatively weighted harmonics which break the congruence of phase at the zero-cross point. However, the periodic halt of the phase advance of the harmonics causes the phases to pile up around ±90° at the points where the ramps meet plateaux. A phase grouping around 90° is typical of that produced by delta functions and bars and may be the signal that produces Mach bands.

Why should the visual system react so strongly to phase relationships? A two dimensional basis is required to extract phase information. The even and odd symmetry of receptive fields of the visual cortex[11] could form such a basis. The receptive fields vary in size, also allowing for coding of an image at different spatial scales[12-15]. The potential ability of such a system to signal edges (odd-symmetric) and lines (even symmetric) has been noted by many investigators[15-18]. We suggest that an efficient means for the visual system to locate bars and edges would be to consider the sum of the squared output of even and odd symmetric filters, which always peaks at points of phase congruence. After the peak had been located the response of the odd and even fields at that point would determine if it were due to an edge or a bar, and give the sign and contrast of the edge or bar. The clustering of phase around 90°, which occurs with trapezoids, should also signal a bar or stripe.

Phase is thought to be encoded by even and odd symmetric detectors in the human visual system[19]. It has further been shown that the sensitivity of the odd, but not the even, symmetric mechanism is reduced in peripheral vision[20]. This would account for the disappearance of Mach bands in the periphery, at contrasts well above those required for the independent detection of out-of-phase harmonics.

Although the generally accepted explanation of Mach bands is that they are due to lateral inhibition, several investigators have noted the problem this presents for square waves and sharp edges[2,5,6,16] and have considered other possibilities. Tolhurst observed that even symmetric receptive fields (which he termed bar detectors) may signal Mach bands[16]. He also speculated that for the square wave the response of odd symmetric fields (edge detectors) may inhibit that of the even symmetric fields (bar detectors), so no stripes are seen (see also ref. 6). Watt and Morgan's[21] general theory of spatial vision also predicts bar signals at the border of a ramp, provided that they are far enough apart. Both these ideas have some similarity to ours, but our results (and simulations to be reported in a fuller paper) suggest that inhibition between edge detectors and bar detectors is unnecessary.

Finally, we note that trapezoids, like square waves and synthetic 'all positive' trapezoids, show zero-crossings at all scales at the mean luminance point. If edges were encoded by alignments of zero-crossings, as has been assumed[22-24], all should have edges at that point. But trapezoids do not. A model seeking phase congruence as the signature of bars and edges, built on adequate basis functions, can without ambiguity, locate edges and bars where we see them.

Received 2 June; accepted 2 October 1986.

1. Ratliff, F. *Mach Bands* (Holden-Day, San Francisco, 1965).
2. von Bekesy, G. *Vis. Res.* **8,** 1483-1499 (1968).
3. Cornsweet, T. N. *Visual Perception* (Academic, New York, 1970).
4. Fiorentini, A. in *Handbook of Sensory Physiology* VII/4, 188-201 (Springer, Berlin, 1972).
5. Ross, J., Holt, J. J. & Johnstone, J. R. *Vis. Res.* **21,** 1165-1167 (1981).
6. Ratliff, F. *Vis. Res.* **24,** 163-166 (1984).
7. Mach, E. *I.S.-B. Akad. Wiss. Wien. math.-math. nat. Kl* **54,** 303-322 (1865) (translated in ref. 1).
8. Mach, E. *I.S.-B. Akad. Wiss. Wien. math.-nat. Kl* **115,** 633-648 (1906) (translated in ref. 1).
9. Rentschler, I. & Treutwein, B. *Nature* **313,** 308-310 (1985).
10. Braddick, O. J. *Doc. Opthal. Proc. Series* **30,** 255-262 (1981).
11. Hubel, D. H. & Wiesel, T. N. *J. Physiol.* **160,** 106-154 (1962).
12. Maffei, L. & Fiorentini, A. *Vis. Res.* **17,** 257-264 (1977).
13. De Valois, R. L., Albrecht, D. G. & Thorell, L. G. *Vis. Res.* **22,** 545-559 (1982).
14. Maffei, L., Morrone, M. C., Pirchio, M. & Sandini, G. J. *Physiology* **296,** 27-47 (1979).
15. Kulikowski, J. I. & Bishop, P. O. *Expl. Brain Res.* **44,** 386-400 (1981).
16. Tolhurst, D. J. *Vis. Res.* **12,** 797-804 (1972).
17. Pollen, D. A. & Ronner, S. F. *Science* **212,** 1409-1411 (1981).
18. Pollen, D. A. & Ronner, S. F. *Vis. Res.* **22,** 101-118 (1982).
19. Field, D. J. & Nachmias, J. *Vis. Res.* **24,** 333-340 (1984).
20. Bennett, P. & Banks, M. *Nature* (submitted).
21. Watt, R. J. & Morgan, M. J. *Vis. Res.* **24,** 1387-1397 (1984).
22. Marr, D. *Vision* (Freeman, San Francisco, 1980).
23. Marr, D. & Hildreth, E. *Proc. R. Soc. Lond.* **B207,** 187-217 (1980).
24. Hildreth, E. C. *MIT. A.I. Lab. Memo* **858,** (1985).

Hysteresis in the perception of motion direction as evidence for neural cooperativity

Douglas Williams, Gregory Phillips & Robert Sekuler

Departments of Neurobiology/Physiology and Psychology,
Northwestern University, Evanston, Illinois 60201, USA

When elements of a parallel network, such as the human brain, are extensively interconnected, the network can exhibit 'cooperative behaviour'. Such behaviour, which is characterized by order–disorder transitions, multi-stable states, and a form of memory called 'hysteresis', has been observed in human stereopsis[1,2] and has motivated models of stereopsis that incorporate cooperative networks[3–6]. More recently, cooperative phenomena have also been observed in human visual motion perception[7]. This report strongly supports a cooperative interpretation of motion perception by demonstrating hysteresis in the perception of motion direction. The results agree quantitatively with a mathematical model incorporating nonlinear excitatory and inhibitory interactions among direction-selective elements.

Our stimuli were dynamic random dot cinematograms comprising 512 computer-generated dots. Each dot took an independent, two-dimensional random walk of constant step size (0.9°). The direction in which any dot moved was independent of its own previous displacements and also of the displacements of the other dots[8].

The direction of motion for each dot was chosen from either (1) one of two uniform distributions, or (2) a mixture of these two distributions. When all dots drew their movements from a uniform distribution extending over 360°, only the 'local random motion' of individual dots was evident. This resembled a detuned television set. However, when all dots drew their movements from a uniform distribution of 180° or less, the dots appeared to flow *en masse* in the direction of the distribution mean—although individual perturbations were still evident. This resembled snowflakes that, although individually perturbed by wind currents, seem to drift in one direction. In all experiments the mean of the signal distribution was upward.

For convenience, we refer to the 180° and 360° distributions as the signal and noise distributions, respectively. When dots drew their movements from a combination of these distributions, each dot's displacement came randomly from either distribution. The resulting perception depended upon the relative proportions of noise and signal.

Two types of trials were run in random order. In one, the direction of motion for each dot was chosen initially from the signal distribution, producing a perception of upward flow. After a random interval lasting up to 12 seconds, the proportion of dots drawing directions from the signal distribution was decreased and the proportion of dots drawing directions from the noise distribution was increased. The observer responded when the cinematogram changed in appearance from upward flow to local random motion. After this response, the proportion of signal continued to decrease for a random interval (up to 6 seconds). The process was then reversed, with the proportion of signal increasing until the subject responded that the upward flow had reappeared. This terminated the trial. For the second trial type, the stimulus sequence was reversed. We started with an initial percept of local random motion, and measured the points of transition to a perception of upward flow and back again to local random motion. These two different trial structures—signal-first and noise-first—were designed to produce different histories of directional exposure and thus reveal any perceptual effects dependent upon the history of stimulation. For both trial types, the signal to noise ratio changed slowly, with the proportion of dots sampling from each distribution changing by just two dots per frame. At the frame rate of our display, 10 Hz, it took a minimum of 25 seconds for the display to shift from complete dependence on one distribution to complete dependence on the other.

Observers used one eye to watch the centre of a circular display 16° in diameter. The dots were presented at twice threshold luminance against a dim background. Over the course of 5 sessions we made 100 measurements of the signal proportion at the two perceptual transitions for each of our three naive observers.

We repeated this experiment with two narrower ranges of signal distribution: 90° and 1°. Each of the three signal distributions generated a different history of exposure. We expected that, as the distribution narrowed, the signal would become

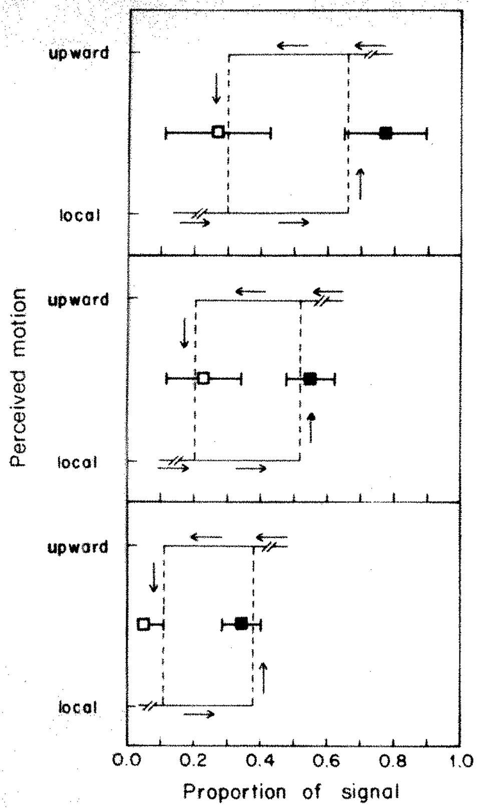


Fig. 1 Perceptual transitions measured under two different histories of stimulus exposure (indicated by arrows) for three signal distributions. The results, for observer J.F., were obtained for signal distributions whose ranges were: top, 180°; middle, 90°; bottom, 1°. Data points show proportion of 'signal' dots required for perceptual transition from local random motion to upward flow (■) and for perceptual transition from upward flow to local random motion (□). Error bars indicate one standard deviation (100 measurements). In each panel the separation between transition points measured with the different exposure histories is an index of hysteresis. The dashed lines mark the transition points calculated from a model incorporating cooperative interactions among direction-selective motion elements.

more effective in stimulating directionally selective visual elements that are tuned to upward motion. As a consequence, the occurrence of a transition between perceived states would require fewer signal dots.

The results for one observer are shown in Fig. 1; similar results were obtained for the other observers. Each panel shows data for a different range of signal. Note that the transitions differ for the three signal ranges. More importantly, the transition from noise-to-upward flow requires a larger number of dots sampled from the signal distribution than does the transition from upward flow-to-noise. For all three signal distributions, the two types of transition occur at significantly different ratios of signal to noise ($P<0.005$).

Before attributing our results solely to neural hysteresis, a number of alternative explanations were considered and rejected. For example the motion after-effect, or waterfall illusion, did not produce our results as this after-effect would probably have facilitated, rather than retarded, the transition from upward motion to noise. Also, the results were largely unchanged by the use of a fixation point, suggesting that eye movements probably played little or no role. Finally, reaction time could be dismissed given the slow time course for changing the signal proportion.

Before developing a simple network that could describe our results, we needed a fuller account of how spatial parameters might affect hysteresis. Using a 180° signal distribution, we made measurements (1) with a four-fold decrease in the spatial density of the cinematogram's dots, (2) with a four-fold decrease in the area of the display and (3) with the display shifted horizontally into the periphery of the visual field so that the nearest dot was 4° from fixation. None of these variations produced results that differed greatly from our original measurements. Of course, extreme changes in these display parameters would undoubtedly alter the hysteresis characteristics. However, because of these data and demonstrations suggesting that spatial variables have little effect on the dot interactions responsible for motion in cinematograms[8,9], we chose not to treat space explicitly in developing a model network to describe our results.

The model comprises a set of direction-selective mechanisms covering all 360° of motion direction, with each direction-selective mechanism having a gaussian profile for directional sensitivity. Based on previous results from our laboratory, the half-amplitude half-bandwidth of each mechanism's gaussian sensitivity profile was set to 30° (ref. 10). The model, whose mathematical formulation is a modification of the cooperative neural network previously proposed[4], assumes non-linear excitatory interactions among mechanisms sensitive to similar directions of motion and non-linear inhibition among mechanisms sensitive to different directions. The dynamic response of this cooperative system is represented by a pair of coupled differential equations. For the excitatory activity, E_i, in direction channel i:

$$\mathrm{d}E_i/\mathrm{d}t = -E_i + (1-E_i)S\left(P_i + \sum_j a_j E_j - \sum_j b_j I_j\right)$$

where S is a nonlinear function of sigmoidal shape, P_i is the external input to channel i, and a_j, b_j are the excitatory and inhibitory weights, respectively, of channel j with respect to channel i. A similar equation gives the inhibitory activity I_i, in channel i. In general terms, such interactions promote the formation of stable 'coalitions' among similarly tuned elements within the network[11]. These neural coalitions can in turn produce various cooperative properties, including hysteresis.

The parameters of the model were constrained so that the model behaved in what is defined[4] as the active transient mode. In this mode the system shows hysteresis, switching between different states of activity. In the model simulation the perception of local random motion is represented by a steady state of uniform activity across all mechanisms. Global upward flow is represented by a steady state in which the activity is localized about the mechanism selective for upward movement. A transition point is defined by the proportion of signal at which the network switches between these two states of activity. The results from the model are shown in Fig. 1. Dashed lines mark the transition points calculated from the model using a single parameter set. It can be seen that the model captures both the leftward shift and narrowing of the hysteresis profile with decreasing signal range. The model's behaviour here is readily explained. With decreasing signal range more activity is concentrated in fewer motion-selective elements arrayed about the upward direction. Consequently, a smaller proportion of signal dots suffices to indicate upward movement. In addition, fewer active elements reduces the opportunity for cooperative interactions in the network, which translates into a narrowing of the hysteresis profile[2].

If we accept the idea that the perception of motion direction depends upon a cooperative neural network, there may be circumstances outside the laboratory in which this network's cooperativity might be especially useful. Observers must extract a mean direction vector from a scene containing a great many different local vectors in many naturally-occurring situations. For example, we can judge the average direction in which the ocean's surf moves despite the fact that individual waves move along somewhat different paths; or we can judge the average direction in which the wind blows the leaves of a tree despite the random variation of that movement from one leaf to the next (or for any one leaf over time). Faced with such multivectorial stimuli, a cooperative network like the one embodied in our model would enhance the signal-to-noise ratio and thereby facilitate the perception of the mean direction of motion.

Supported by a grant from the US Air Force Office of Scientific Research.

Received 12 May; accepted 2 September 1986.

1. Julesz, B. *Am. Sci.* **62,** 32-43 (1974).
2. Fender, D. H. & Julesz, B. *J. Opt. Soc. Am.* **57,** 819-830 (1967).
3. Sperling, G. *Am. J. Psychol.* **83,** 461-534 (1970).
4. Wilson, H. R. & Cowan, J. *Kybernetik* **13,** 55-80 (1973).
5. Marr, D. & Poggio, T. *Science* **194,** 283-287 (1976).
6. Mayhew, J. E. W., Frisby, J. P. & Gale, P. *Perception* **6,** 207-208 (1977).
7. Chang, J. J. & Julesz, B. *Vis. Res.* **24,** 1781-1788 (1984).
8. Williams, D. W. & Sekuler, R. *Vis. Res.* **24,** 55-62 (1984).
9. Baker, C. L., Jr. & Braddick, O. J. *Vis. Res.* **22,** 1253-1259 (1982).
10. Williams, D. W., Tweten, S. & Sekuler, R. *Invest. Ophthalmol. vis. Sci. Suppl.* **25,** 14 (1984).
11. Feldman, J. A. & Ballard, D. H. in *Human and Machine Vision* (eds Beck, J., Hope, B. & Rosenfeld, A.) 107-155 (Academic, New York, 1983).

Inhibitory neurones of a motor pattern generator in *Xenopus* revealed by antibodies to glycine

N. Dale*†, O. P. Ottersen‡, Alan Roberts§ & J. Storm-Mathisen‡||

‡ Anatomical Institute, University of Oslo, Karl Johans gate 47,
N-0162 Oslo 1, Norway
* Department of Physiology III, Karolinska Institute,
Lidingövägen 1, Stockholm S-114 33, Sweden
§ Department of Zoology, University of Bristol, Woodland Road,
Bristol BS8 1UG, UK

Glycine and γ-aminobutyric acid (GABA) are inhibitory transmitters of major importance[1,2]. Whereas neurones using GABA as the transmitter can be visualized by immunocytochemical methods for glutamate decarboxylase (GAD)[3,4] or GABA[5-8], no comparable techniques have been available for the selective visualization of glycinergic neurones. We have now produced polyclonal antibodies which specifically recognize glycine in glutaraldehyde-fixed tissue. We used these antibodies to investigate the distribution of glycine in the simple central nervous system (CNS) of the *Xenopus* embryo, which contains an anatomically and physiologically defined class of reciprocal inhibitory interneurones, the commissural interneurones[9-11]. These interneurones have an important role in the generation of the swimming motor pattern and are thought to be glycinergic[11-13]. The glycine antibodies specifically stain these interneurones, revealing their distribution and number in the embryo CNS. This is the first demonstration of the selective localization of glycine-like immunoreactivity in a putative glycinergic class of neurone that has been characterized physiologically, pharmacologically and anatomically.

Evidence for glycine as a transmitter in the vertebrate spinal cord has been based on four types of information: the actions of exogenous glycine[14,15], specific blockade by the antagonist strychnine[14,15], the depletion of glycine following ischaemic degeneration of spinal interneurones[16,17] and the demonstration of neuronal uptake of ^{3}H-glycine[18-21]. However, antagonism of an inhibitory postsynaptic potential (i.p.s.p.) by strychnine is not unequivocal evidence that glycine is the endogenous transmitter: the hyperpolarizing actions of β-alanine and taurine (amino acids closely related to glycine; Fig. 1) are also blocked by strychnine[2]. In the absence of further evidence β-alanine and taurine must therefore also be considered as candidate transmitters. We decided to raise antibodies which could selectively recognize glycine in glutaraldehyde-fixed nervous tissue and could be used to localize glycine in neurones. Modifications of our methods[5,6] have been used previously by Pourcho and Goebel[22] to produce and test a glycine antiserum which marks cells in the cat retina and brain stem.

Glycine was coupled to carrier proteins by means of glutaraldehyde as described previously[5]. As suggested by Seguela *et al.*[7], the carrier protein was changed, here for each immunization, to avoid boosting the immune response to the carrier protein and any particular configurations of the glycine-glutaraldehyde complex that might depend on the protein (Fig. 1 legend). Unwanted antibodies recognizing glutaraldehyde-treated amino acids and proteins were removed by immunosorbent chromatography on a sequence of columns (Fig. 1 legend). The purified antiserum was initially tested for crossreactivity with a series of compounds conjugated by glutaraldehyde to a total macromolecular extract from rat brain and spotted on cellulose ester filters[6]. The compounds were GABA, taurine, phosphoethanolamine, β-alanine, L-α-alanine, valine, leucine, proline, serine, threonine, tryptophan, tyrosine, histidine, lysine, arginine, ornithine, aspartate, *N*-acetylaspartate, glutamate, asparagine, glutamine, methionine, cysteine, reduced glutathione, α-aspartylglycine, carnosine, homocarnosine, cadaverine, putrescine, spermine and spermidine. While the glycine conjugate was intensely stained, none of the other conjugates were stained appreciably darker than 'null' conjugates (macromolecular extract treated with glutaraldehyde in the absence of amino acid). The staining was abolished by solid-phase absorption with glycine fixed to ovalbumin or to lysine by glutaraldehyde (Fig. 1 legend).

To check the specificity of these antibodies in the *Xenopus* embryo we made two different sets of conjugates, one of macromolecules from whole embryos and one of macromolecules from the CNS of adult *Xenopus*. To ensure identical conditions for testing and immunocytochemistry, the test filters with these conjugates were processed in the same vessels as the tissue preparations. The tissue consisted of the CNS of *Xenopus* embryos stage 29/30 to 37/38 (ref. 23) raised by induced breeding in Bristol, fixed in glutaraldehyde and shipped to Oslo for dissection and processing[6] of the CNS.

The results with *Xenopus* conjugates (Fig. 1*a*) were similar to those with rat brain conjugates. A glycine-glutaraldehyde complex, added to the diluted antiserum (at a final concentration of 100 or 300 μM with respect to glycine) completely abolished staining of the glycine spots and of the tissue (Fig. 1*e*, *f*). Complexes of β-alanine (Fig. 1*c*, *d*), L-α-alanine, GABA or reduced glutathione (all at 300 μM), or a mixture of individual glutaraldehyde complexes of taurine, aspartate, glutamate, glutamine and serine (all at 300 μM) produced no reduction in the staining of glycine test spots (the converse is seen in ref. 32) or tissue.

The *Xenopus* embryo spinal cord appears on the basis of retrograde horseradish peroxidase (HRP) filling to contain eight anatomically distinct neurone classes[9]. One of these, the commissural interneurones, is defined by the following anatomical features[9]: (1) a unipolar soma in a middle to dorsal position in the cord, (2) a single ventral process bearing dendrites that extend laterally into the marginal zone (lateral tract) and (3) an axon crossing ventrally to ascend and sometimes also descend contralaterally (Fig. 3*a*). These neurones have been shown by intracellular recording and HRP injection to be active during fictive swimming[10] and to produce mid-cycle i.p.s.ps in contralateral motoneurones[11]. The mid-cycle i.p.s.ps in motoneurones and the response to exogenous glycine (100 μM) have a reversal level near the resting potential, depend on an increase in chloride permeability and are unaffected by 10 μM bicuculline, but are blocked by 1 μM strychnine[11,13,24]. The commissural interneurones may therefore be glycinergic.

---

† Present address: Howard Hughes Medical Institute, Center for Neurobiology and Behavior, Columbia University, 722 West 168th Street, New York 10032, USA.

|| To whom correspondence should be addressed.

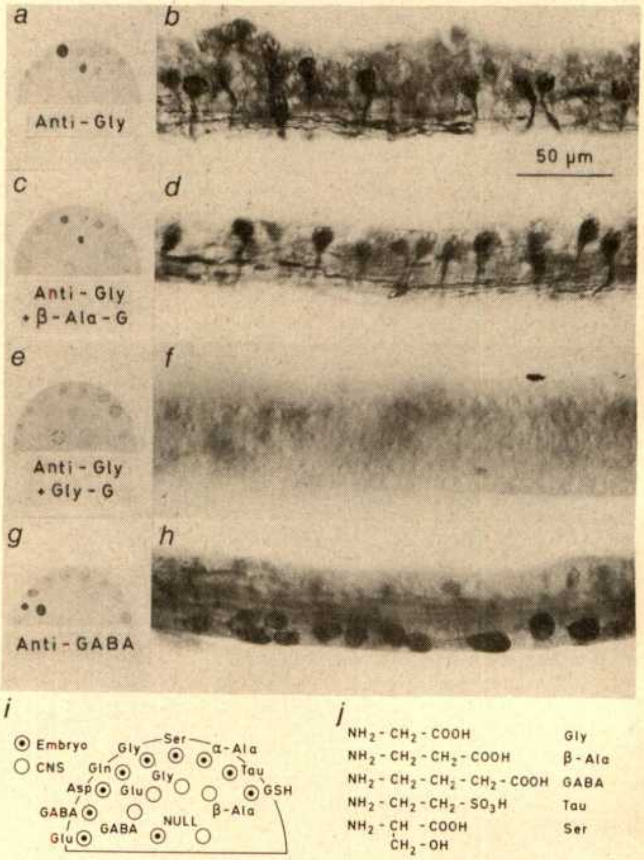


Fig. 1 The purified glycine antiserum (31, 1:90) gives selective staining of test spots of glycine fixed by glutaraldehyde to macromolecules from *Xenopus* embryos (pigmented) (⊙) or adult CNS (○) (*a*, *i*), and selective staining of a characteristic class of neurones, the commissural interneurones, in whole-mounts of the spinal cord of *Xenopus* embryos (lateral view) (*b*). The specific staining is abolished by glutaraldehyde complexes of glycine (Gly-G) (*e*, *f*), but not by complexes of the closely related β-alanine (300 μM added to the diluted serum) (*c*, *d*). In comparison, a purified GABA antiserum (26, 1:200)[6] stains only the GABA spots (*g*) and different cells[25] (*h*) from those stained by serum 31. The structures of the amino acids most closely related to glycine are shown (*j*); standard symbols are used for amino acids; GSH, γ-glutamylcysteinylglycine (glutathione).

Methods. The serum was raised by immunizing a rabbit (intracutaneously on the back at 6-week intervals with 125 μg protein in Freund's complete adjuvant; Gibco) with glutaraldehyde conjugates (0.3–1 μmol of glycine bound per mg protein) of a sequence of carrier proteins: keyhole limpet haemocyanin (Calbiochem), casein (Sigma), human haemoglobin (Sigma), lysozyme (Worthington), *Limulus polyphemus* haemocyanin (Sigma), horse skeletal muscle myoglobin (Sigma) and bovine thyroglobulin (Sigma). Portions (50 μl) of the crude serum were applied to a set of three microcolumns (i.d. 0.3 cm) connected in series and containing protein (6 mg per ml gel) coupled to agarose particles (Sepharose 4B, Pharmacia[27]) and reacted with glutaraldehyde (Taab, EM-grade, 75 μmol per ml gel) in the presence or absence of amino acid (20 μmol per ml gel). Column 1 carried bovine serum albumin treated with glutaraldehyde, column 2 carried albumin treated with glutaraldehyde in the presence of glutamate, and column 3 consisted of portions of agarose particles coupled with the individual carrier proteins and treated with glutaraldehyde. The serial column was eluted with Tris buffer (0.1 M *p*H 7.4) containing NaCl (0.3 M). Glycine buffer (0.1 M *p*H 2.3 with 0.3 M NaCl) or propionic acid (2 M) was then used to regenerate the column. Columns 1 and 2 were essential for specificity, column 3 did not change the specificity of staining of test conjugates or tissue, but slightly reduced the background. Absorption of the purified serum with mycobacteria from the adjuvant did not change the staining of test conjugates or tissue. The specific antibodies could be completely removed by passage through a microcolumn of agarose particles carrying glycine fixed to ovalbumin by glutaraldehyde, or preformed complexes of glycine and glutaraldehyde (see below) attached to lysine-Sepharose (Pharmacia). CNS of *Xenopus* embryos (fixed in 5% glutaraldehyde diluted from 25% in 0.1 M sodium phosphate buffer *p*H 7.4 (NaP_i)) and test filters[6] were incubated together in the purified serum (dilution 1:90–1:400, 3–6 days, 4–6 °C) with 0.5% Triton X-100 and further processed essentially as described previously[6]. Penetration of reactants was enhanced by soaking the CNS in sucrose (30% in NaP_i), freezing in liquid N_2 and thawing at room temperature before incubation with the antisera. The test conjugates spotted on the filter semicircles were prepared by mixing 4.5 mg tissue protein per ml in NaP_i with 20 mM amino acid and 50 mM glutaraldehyde for 1 h, and dialysing. Specificity was further demonstrated by adding glutaraldehyde complexes of various amino acids to the diluted serum. The complexes were prepared by mixing amino acid (20 μmol) with glutaraldehyde (40 μmol, that is, optimal stoichiometry[28], total volume 116 μl) for 1 h at room temperature and adding ethanolamine HCl (1 M 384 μl), all in NaP_i (see also ref. 31).

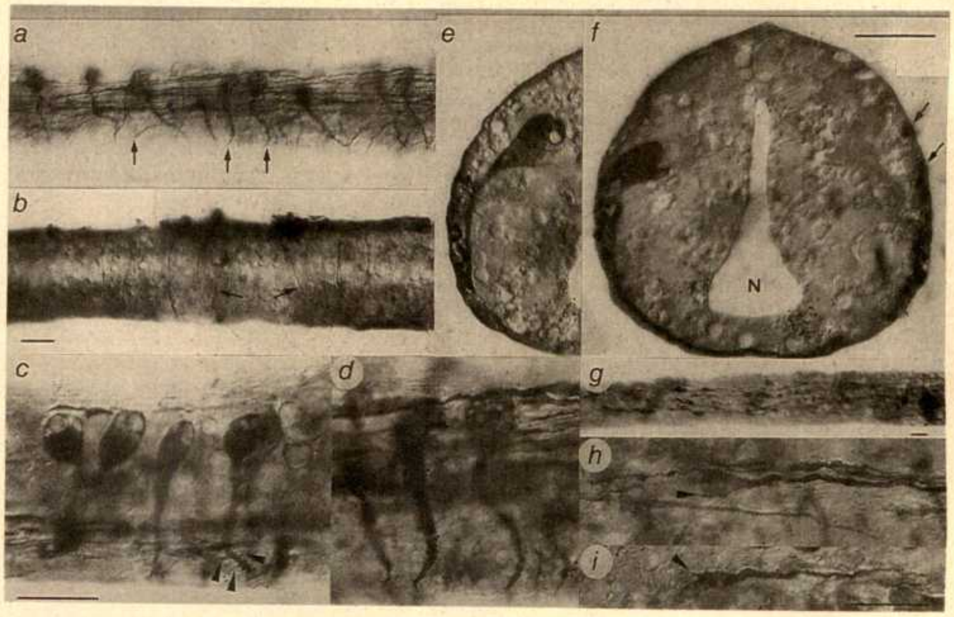


Fig. 2 Characteristic anatomical features of the commissural interneurones of *Xenopus* embryo spinal cord (see also Fig. 3) revealed by specific staining with the glycine antiserum (see Fig. 1 legend). *a*, Lateral view of a slightly rotated spinal cord showing unipolar somata with axons (arrows) crossing ventrally under the neurocoel. *b*, Ventral view of the spinal cord showing the crossing axons (arrows). *c*, *d*, High-power lateral view of stained cells with dendrites (arrowheads) coming out of the fat proximal process. *e*, *f*, Transverse sections (6 μm) of immunostained spinal cords osmicated and embedded in Araldite showing two unipolar somata with axons crossing ventrally under the neurocoel (N). The marginal zones clearly contain transversely cut stained axons (arrows). *g*–*i*, Stained axons and growth cones in the lateral axon tracts. Arrowheads show short micropodia emanating from growth cones. The large unstained inclusions are yolk platelets which are present in all the embryonic cells. Interference contrast optics; all from whole-mounts except *e*, *f*. Scale bars, 20 μm.

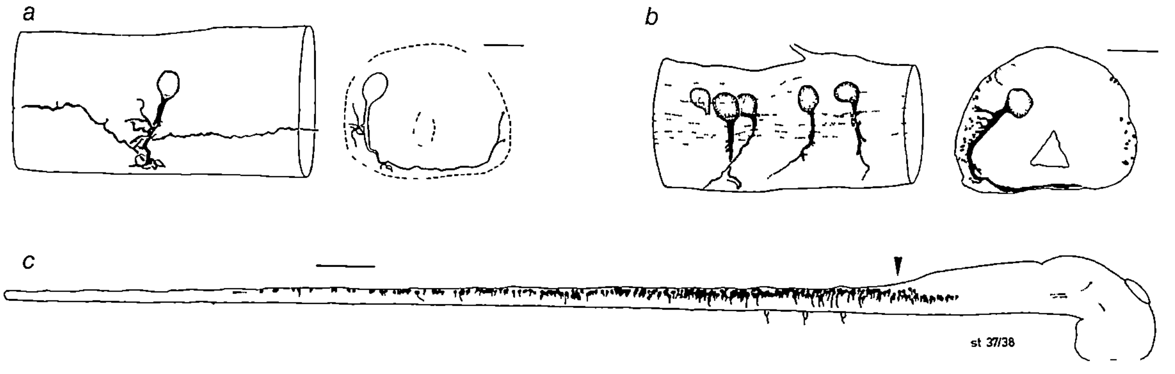


Fig 3 Camera lucida drawings of commissural interneurones from *Xenopus* embryo spinal cord *a*, Commissural interneurone filled with horseradish peroxidase from the intracellular recording micropipette in a stage 37/38 embryo Left, lateral view, right, reconstruction of the transverse section view of these cells (based on micrometer readings in the side view) showing dendrites and axon passing under cord before forming ascending and descending branches (redrawn from ref 10) *b*, Neurones stained with the glycine antiserum showing similarity with *a* Left, drawing of trunk spinal cord from whole-mount at stage 33/34 (lateral view) Four of the stained cells have a unipolar soma with their main ventral process with short dendrites, and an axon which runs ventrally to cross to the other side Dotted lines show stained axons Right, 6-μm transverse plastic section of stage 37/38 rostral trunk cord Stained axons are present throughout the marginal zone on each side One faint and one clear soma are present on the left side The latter has a large ventral process giving off dendrites into marginal zone and continuing ventrally to run under neurocoel to the opposite side *c*, Selective distribution of glycine-positive neurones along the neuraxis in a stage 37/38 embryo Arrowhead shows border between hindbrain and spinal cord All neurones (299) on one side are shown Scale bars, *a*, *b*, 25 μm, *c*, 200 μm

To investigate whether these neurones contain an elevated glycine concentration, we processed the embryo CNS with the glycine antibodies This gave specific staining of a population of cells with the three defining features of commissural interneurones outlined above (Figs 2*a*–*f*, 3*b*) At stage 37/38[23] there were about 300 cells per side extending from the mid-hindbrain to the caudal spinal cord (Fig 3*c*) Comparison with results from retrograde HRP filling indicated that the glycine antiserum had stained the whole population of commissural interneurones The most caudal somata were often paler than the rest, suggesting that cells in this region were still differentiating A few of the stained somata lay close to the cord surface in a dorsolateral position (Fig 3*b*) These cells may either be commissural interneurones with slightly different anatomy, or members of the other class of spinal interneurone with crossed axons, the dorsolateral commissural interneurones[9] Staining of ventral commissural and longitudinal axons was darker than that in somata Some axons bifurcated before reaching the other side Ascending axons reached the rostral hindbrain while a few extended into the midbrain or even the forebrain Descending axons reached to within 600 μm of the caudal end of the spinal cord at stage 37/38 Many axons terminated in large, stained growth cones (Fig 2*g*–*i*) No other groups of somata were stained

An antiserum specific for GABA[6] stained spinal interneurones with ascending axons and one class of possible receptor cells, the ciliated ependymal cell (Fig 1*h*)[9,25] However, it did not stain any commissural interneurones or ventral commissural axons in the spinal cord Antisera specific for taurine[26] or β-alanine (S Davanger, A Bruun, B Ehinger, J Laake, L Orensanz, O P O & J S -M , in preparation) produced no staining of cells or fibres in the *Xenopus* embryo spinal cord Our immunocytochemical data together with the pharmacological and physiological evidence thus suggest that glycine rather than another inhibitory amino acid is the transmitter of the commissural interneurones

We thank A T Bore, T Eliassen, I Fridstrøm, G F Lothe, K Stentoft, L J Teagle and J Line Vaaland for assistance This work was supported by the Norwegian Research Council for Science and the Humanities, the SERC, the Company of Biologists, the Norwegian Council on Cardiovascular Disease and the NFtKB

Note added in proof Campistron *et al*[29] have recently reported on an antiserum to glutaraldehyde-fixed glycine (reacted with $NaBH_4$) and used it to stain neurones in rat brain and spinal cord With a few exceptions, their results are similar to those obtained in rat with our antiserum[30]

Received 30 May, accepted 28 July 1986

1 Curtis, D R. & Johnston, G A. R. *Ergebn. Physiol* **69**, 97–188 (1974)
2 Krnjević, K. *Physiol Rev* **54**, 418–540 (1974)
3 Saito, K. *et al. Proc. natn. Acad Sci. U S.A.* **71**, 269–273 (1974)
4 Mugnaini, E & Oertel, W *Handbook of Chemical Neuroanatomy* Vol 4 (eds Bjorklund, A & Hökfelt, T) (Elsevier, Amsterdam, 1985)
5 Storm-Mathisen, J *et al. Nature* **301**, 517–520 (1983)
6 Ottersen, O P & Storm-Mathisen, J *J comp Neurol* **229**, 374–392 (1984)
7 Seguela, P , Geffard, M , Buijs, R. M & Le Moal, M *Proc. natn. Acad Sci. U S A.* **81**, 3888–3892 (1984)
8 Hodgson, A., Penke, B , Erdei, A , Chubb, I W & Somogyi, P *J Histochem. Cytochem.* **33**, 229–239 (1985)
9 Roberts, A. & Clarke, J D W *Phil. Trans R. Soc.* **B296**, 195–212 (1982)
10 Soffe, S R., Clarke, J D W & Roberts, A *J Neurophysiol* **51**, 1256–1267 (1984)
11 Dale, N *J Physiol., Lond* **363**, 61–70 (1985)
12 Roberts, A , Soffe, S R & Dale, N in *Neurobiology of Vertebrate Locomotion* (eds Grillner, S *et al.*) (Macmillan London 1986)
13 Soffe, S R *J Physiol., Lond* **382** (1987)
14 Werman, R., Davidoff, R. A & Aprison, M H *Nature* **214**, 681–683 (1967)
15 Curtis, D R , Hosli, L & Johnston, G A R *Expl Brain Res* **6**, 1–18 (1968)
16 Davidoff, R. A , Graham, L. T Jr, Shank, R. P , Werman, R. & Aprison, M H *J Neurochem.* **14**, 1025–1031 (1967)
17 Homma, S , Suzuki, T , Murayama, S & Otsuka, M *J Neurochem.* **32**, 691–698 (1979)
18 Hökfelt, T & Ljungdahl, Å. *Brain Res* **32**, 189–194 (1971)
19 Ljungdahl, Å. & Hokfelt, T *Brain Res* **62**, 587–595 (1973)
20 Matus, A I & Dennison, M E *Brain Res* **32**, 195–197 (1971)
21 Price, D L , Stocks, A., Griffin, J W , Young, A. & Peck, K. *J Cell Biol* **68**, 389–395 (1976)
22 Pourcho, R G & Goebel, D J *Brain Res* **348**, 339–342 (1985)
23 Nieuwkoop, P D & Faber, J *Normal Tables of Xenopus laevis* (*Daudin*) (North-Holland, Amsterdam, 1956)
24 Soffe, S R & Roberts A *J Neurophysiol* **48**, 1279–1288 (1982)
25 Dale, N , Roberts, A., Ottersen, O P & Storm-Mathisen, J (in preparation)
26 Madsen, S , Ottersen, O P & Storm-Mathisen, J *Neurosci. Lett* **60**, 255–260 (1985)
27 Porath, J , Aspberg, K., Drevin, H & Axén, R. *J Chromatogr* **86**, 53–56 (1973)
28 Hardy, P M , Nicholls, A C & Rydon, H N *J C S Perkin Trans* **9**, 958–962 (1976)
29 Campistron G Buijs R M & Geffard M *Brain Res* **376**, 400–405 (1986)
30 Ottersen, O P , Davanger, S & Storm Mathisen, J *expl Brain Res* (in the press)
31 Ottersen, O P , Storm-Mathisen, J , Madsen, S , Skumlien, S & Strømhaug, J *med. Biol* **64** 147–158 (1986)

T-cell epitope of the autoantigen myelin basic protein that induces encephalomyelitis

Scott S. Zamvil*, Dennis J. Mitchell*, Anne C. Moore*, Kumiko Kitamura*, Lawrence Steinman* & Jonathan B. Rothbard†‡

* Departments of Neurology and Pediatrics, Stanford University, Stanford, California 94305, USA
† The Imperial Cancer Research Fund, P O Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

Chronic relapsing paralysis and demyelination within the central nervous system (CNS), features associated with the human disease multiple sclerosis (MS)[1-4], develop in mice after injection of murine T-cell clones specific for the autoantigen myelin basic protein (MBP)[5,6]. We examined the fine specificity of three independently derived encephalitogenic T-cell clones using synthetic polypeptides derived from portions of the N-terminal sequence of MBP. These clones appear functionally identical; they all respond to an epitope in the N-terminal nine amino acid residues in association with the same class II (I-A) molecules of the major histocompatibility complex (MHC). Both the N-terminal acetyl moiety and the first residue (Ala) are necessary for recognition. Only N-terminal MBP peptides recognized by these clones were found to cause encephalomyelitis (EAE) *in vivo*. These results show that the N-terminal MBP-specific T lymphocytes that mediate autoimmune encephalomyelitis are a small population with a limited repertoire; they all recognise the same combination of MHC and target.

EAE is a model for autoimmune diseases mediated by class II restricted T lymphocytes[5-10] Immunization with intact MBP or fragments of MBP causes EAE in susceptible mouse strains[10-12] However, not all such fragments are encephalitogenic Only the N-terminal 37 amino-acid MBP fragment causes EAE in PL/J (H-2^u) and (PL/J × SJL/J (H-2^s))F_1 ((PLSJ)F_1) mice[11,12] Encephalitogenic T-cell clones with the $L3T4^+$, $Lyt2^-$ phenotype have been isolated from both these strains following immunization with either intact rat or bovine MBP[5,6] Clones PJR-25 and F1-12 were isolated from homozygous PL/J and (PLSJ)F_1 mice, respectively, after immunization with intact rat MBP Encephalitogenic clone PJB-20 was isolated from homozygous PL/J mice after immunization with intact bovine MBP These clones respond to a determinant shared with mouse (self) MBP within the encephalitogenic N-terminal MBP fragment in association with I-A^u($A_\alpha{}^u A_\beta{}^u$) class II molecules Class II-restricted T lymphocytes recognize discrete, contiguous sequences within proteins, sometimes fewer than ten residues[13] in length via their antigen-specific cell surface receptors[14,15] To investigate the pathogenesis of this autoimmune disease further, we examined whether these encephalitogenic T lymphocytes all respond to a common epitope, or have distinct N-terminal MBP specificities

Separate forms of intact MBP with different N-terminal sequences were tested for their ability to stimulate proliferation of encephalitogenic T-cell clones Bovine MBP was found to be less stimulatory than mouse or rat MBP at the same concentration (results not shown) Since the N-terminal 37 amino-acids of bovine MBP differ from mouse MBP only at residues 2 and 17 (see Fig 1)[16], we predicted that the epitope(s) recognized by these T-cell clones would include one of these two residues The overlapping peptides pR(rat)1-16 and pR5-20 were made and tested for their ability to stimulate proliferation of these clones The epitope recognized by all three clones is found in the N-terminal 16 residues, peptide pR1-16 is as effective as intact MBP (Fig 2*a*) whereas pR5-16 and pR5-20 do not stimulate proliferative responses These clones also respond to peptides pR1-9 and pR1-11 Since pR1-11 and pR1-16 are equipotent, residues 12-16 do not contribute significantly to the epitope recognized by these T-cell clones

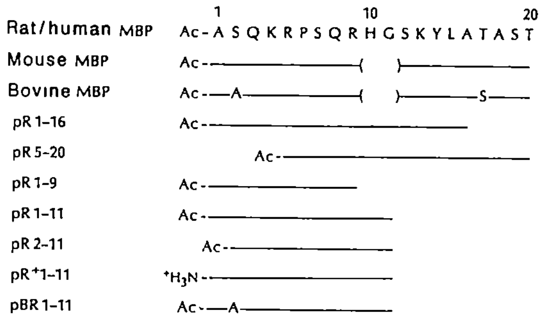


Fig 1 Amino-acid sequences of MBP from different species, and synthetic MBP peptides The sequences of intact rat, human, mouse, and bovine MBP have been published previously[16] Peptides corresponding to the sequences of rat (R) and bovine (B) MBP were synthesized as described previously[29,30] by solid phase techniques on a Beckman Model 990 B peptide synthesizer, using commercially available amino acid polystyrene resins and t-Boc-protected amino acids from Peninsula Laboratories The protecting groups used were o-benzyl esters for Thr and Ser, tosyl for His, orthochlorobenzyloxycarbonyl for Lys, and 2,6-dichlorobenzyl for Tyr Couplings were performed using a 2 5 molar excess of t-Boc amino acid and dichlorohexylcarbodiimide (DCC) over the number of milliequivalents of mino acid on the resin All couplings were >99% complete, as determined by reaction with ninhydrin The peptides were deprotected and removed from the resin simultaneously by treatment with anhydrous hydrogen fluoride in the presence of anisole, dimethylsulphide, and indole, then separated from organic side products by extraction with ether and isolated from the resin by extraction with 5% acetic acid and subsequent lyophilization Purity was determined by high pressure liquid phase column (Merck) and amino acid analysis The peptides used in these experiments were not further purified as they all contained >90% of the desired peptide

Acetylation of the first residue (NAc-Ala_1) is essential for recognition When the acetyl group is removed, revealing a positively charged amino terminus (pR^+1-11, Fig 1), proliferative activity is eliminated (Fig 2*b*) The amino-acid Ala_1 is also necessary, since the MBP peptide pR2-11, containing NAc-Ser_2 is not stimulatory even at the highest concentrations tested Bovine MBP, which has an alanine instead of a serine residue at position 2 (see Fig 1), is less encephalitogenic than intact rat or mouse MBP, which both have Ser_2, in homozygous PL/J and (PLSJ)F_1 mice[11] Peptide pBR1-11, which has Ala as the second residue in a rat MBP1-11, sequence (Fig 1) is less effective than pR1-11 and requires higher concentrations to stimulate proliferation Thus, these representative T-cell clones isolated after stimulation with rat MBP or bovine MBP share the same N-terminal specificity We also identified a second T-cell epitope in the encephalitogenic N-terminal fragment of rat MBP T-cell clones that are restricted to $A\alpha^s A\beta^u$ class II molecules isolated from (PLSJ)F_1 mice recognize rat but not mouse (self) MBP, and are not encephalitogenic The epitope recognized by these T-cell clones is found in pR9-16 (Fig 3) Thus there are two distinct epitopes in the N-terminal 16 residues of MBP

MBP peptide fragments have been shown to cause acute encephalomyelitis in other animal models[17-22] However, peptides as small as those described in this report have not previously been shown to cause EAE in mice Peptides pR1-9, pR1-11, and pR1-16, all of which are recognized *in vitro* by encephalitogenic T cell clones, cause EAE in (PLSJ)F_1 mice when given in adjuvant, without carrier protein (Table 1) Histo-

---

‡ To whom correspondence should be addressed

Table 1 Induction of EAE by synthetic N-terminal MBP peptides

Antigen	Dose (nanomoles)	Incidence	Mean severity	Mean day of onset
pR1-9	100 (110 μg)	10/10	4.2 (±0.63)	11.9 (±0.74)
pR1-9	20	4/5	1.8 (±0.96)	12.5 (±1.7)
pR1-11	100 (128 μg)	10/10	4.4 (±1.3)	12.8 (±1.4)
pR1-11	20	3/5	3.7 (±1.2)	12.3 (±1.2)
pR1-16	100 (186 μg)	3/5	1.7 (±0.57)	14.3 (±1.5)
pR1-16	20	3/5	2.7 (±1.2)	13.7 (±1.2)
pR2-11	200 (244 μg)	0/10	—	—
pR2-11	100	0/5	—	—
pR2-11	20	0/5	—	—
pR5-16	100 (144 μg)	0/5	—	—
pR9-16	100 (97 μg)	0/5	—	—
pR11-27	100 (184 μg)	0/5	—	—
pR20-37	100 (216 μg)	0/5	—	—

MBP peptides were tested for their ability to cause EAE following the protocol described previously[11,12] for induction of EAE in mice with intact MBP. Each peptide was dissolved in phosphate buffered saline (PBS) and emulsified with complete Freund's adjuvant (CFA) in a 1:1 mixture of PBS:CFA containing 4 mg ml^{-1} H37Ra (Difco). Recipient (PLSJ)F_1 female mice (Jackson Laboratories, Maine), aged 12–20 weeks, were injected with 0.1 ml emulsion at the base of the tail. On the same day 10^{10} heat killed *Bacillus pertussis* (Michigan Department of Health, Lot 91B) were injected intravenously. Forty-eight hours later 10^{10} *B. pertussis* were injected a second time. Mice were examined daily for signs of EAE by two independent observers. Severity of EAE was graded using a scale previously described[6]: 0, no sign of EAE, 1, decreased tail tone only, 2, mild paraparesis, 3, moderately severe paraparesis, 4, complete paraplegia, 5, moribund. Synthetic peptides pB1-9, pBR1-11, and pR5-20 have also been tested (results not shown). When (PLSJ)F_1 mice are injected with 100 μmoles pB1-9 or pBR1-11, peptides that are recognized by encephalitogenic T cell clones, clinical and histological EAE are observed. Peptide pR5-20 is not encephalitogenic.

logical examination of paralysed mice shows perivascular infiltration of mononuclear cells in the CNS, as observed in EAE caused by immunization with whole MBP[10,11]. In contrast, peptides pR2-11, pR5-16, pR9-16, pR11-27 and pR20-37, which do not stimulate encephalitogenic T-cell clones to proliferate *in vitro*, do not cause clinical or histological signs of EAE in (PLSJ)F_1 mice (Table 1).

T-cell clones recognize peptide sequences in association with a particular class II MHC product. Encephalitogenic T-cell clones isolated after immunization with intact MBP recognize pR1-11 in association with I-A^u class II molecules (Fig. 2*b*). It is possible that encephalitogenic (PLSJ)F_1 T lymphocytes may recognize this epitope in association with other class II molecules. However, immunization with pR1-11 causes EAE in H-2^u strains (PLJ 7/10 mice paralysed and B10.PL 10/10 paralysed), but not H-2^s strains (SJL/J 0/10 paralysed and B10.S 0/10 paralysed), and lymphocytes primed *in vivo* with pR1-11 are inhibited from responding *in vitro* to either pR1-11 or intact MBP by monoclonal antibodies specific for I-A, but not I-E molecules (results not shown). Thus, the encephalitogenic responses to intact MBP and N-terminal MBP peptides in PL/J and (PLSJ)F_1 mice are restricted to I-A^u class II molecules.

We have thus identified the epitope recognized by T-cell clones which mediates autoimmune encephalomyelitis. These independently-derived T-cell clones appear to have functionally identical peptide specificity and class II restriction. Further, the induction of autoimmune encephalomyelitis by *in vivo* administration of N-terminal MBP peptides correlates with T cell recognition *in vitro*, of all N-terminal MBP peptides tested, only those peptides which are recognized by these independently-derived T-cell clones are encephalitogenic. T-cell clones derived after immunization with pBR1-11 or pR1-11 also have the same N-terminal MBP and class II specificity as the clones described here. Clones derived from mice immunized with pBR1-11 respond in a heteroclitic manner, like PJB-20 they respond better to pR1-11 than to pBR1-11, the original immunogen (results not shown). Thus, *in vitro* and *in vivo* results suggest that the encephalitogenic T-cell response to MBP 1-37 is limited to a discrete population of MBP-reactive T cells.

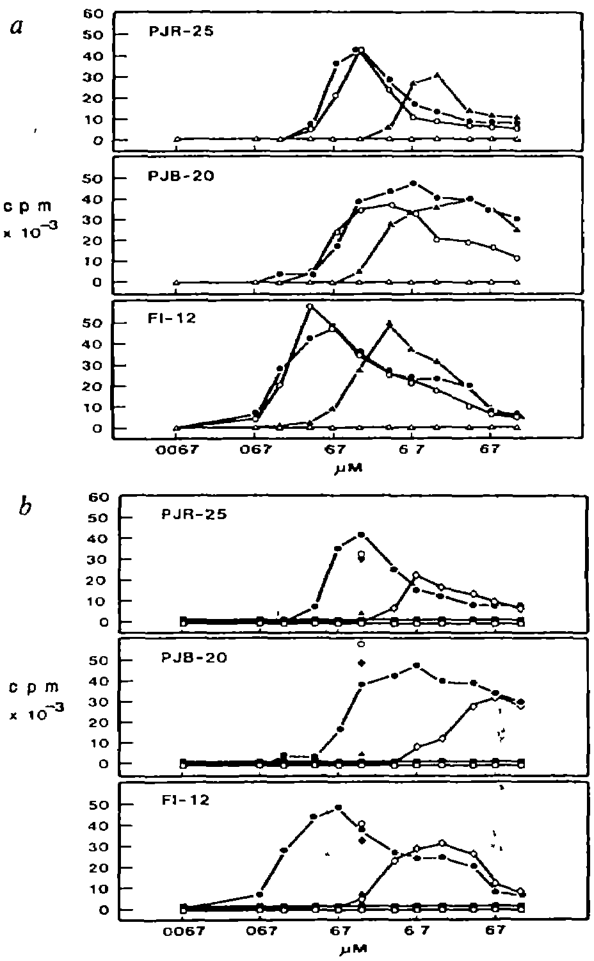


Fig. 2 Recognition of N-terminal MBP peptides by encephalitogenic T cell clones. *a*, Peptides pR1-16 (○), pR1-11 (●), pR1-9 (▲), and pR5-20 (△) were tested for their ability to stimulate encephalitogenic proliferative T cell clones, isolated as described previously[16,31]. To assay for proliferative responses, 10^4 T cells were cultured with 5×10^5 X-irradiated (3,000 rad) PL/J splenic antigen presenting cells (APC) in 0.2 ml culture media in 96 well flat-bottomed microtitre plates (Falcon). Culture medium was RPMI 1640 (Gibco, supplemented with 10% fetal calf serum (Hyclone), 5×10^{-5} M 2-mercaptoethanol, 2 mM glutamine, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin). Peptides were added to cultures giving the final concentrations indicated. At 48 h each well was pulsed with 1 μCi ^{3}H-thymidine and cells collected 16 h later. The mean c.p.m. for thymidine incorporation was calculated for triplicate cultures. Standard deviations are within 10% of the mean. The same pattern of response to N-terminal peptides is found when these T cell clones are cultured with (PLSJ)F_1 APC. *b*, Peptides pR1-11 (●), pBR1-11 (◇), pR^{+}1-11 (□), and pR2-11 (■) were tested as described above. In a separate assay, the proliferative response to pR1-11 (○) was inhibited with 1 μg per well anti I-A monoclonal antibody 10-2.16 (ref. 32) (▲), but not anti I-E monoclonal antibody 14-4-4[33] (◆). When tested at this concentration, 14-4-4 inhibits T cell clones restricted to I-E molecules[6] (results not shown).

It has recently become possible to examine the genes that encode antigen-specific T-cell receptors (TCR)[23–25] and investigate the relationship between specificity and TCR gene expression. Ovalbumin-specific T-cell clones that recognize the same epitope and share the same class II restriction have been shown to use the same V_α/J_α and V_β/J_β TCR genes[26], T-cell clones that share the same fine specificity for cytochrome C use the same V region genes, but can express different J elements[27]. However examination of TCR genes from clones of different specificities suggests that there may not be a simple correlation

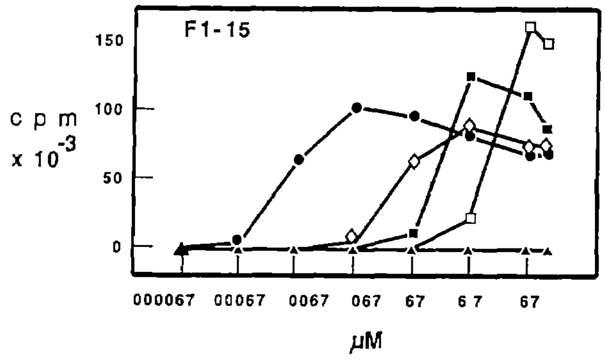


Fig. 3 N-terminal peptides recognized by a representative non-encephalitogenic MBP specific T cell clone Peptides pR1-16 (●), pR5-16 (◇), pR7-16 (■), pR9-16 (□) and pR10-20 (▲), were tested as described (Fig 2, legend) F1-15 T cells were cultured with (PLSJ)F$_1$ splenic APC These N-terminal MBP-specific non-encephalitogenic T cell clones are restricted to $A\alpha^s A\beta^u$ class II molecules N-terminal peptides that are recognized by these non-encephalitogenic T cell clones do not cause EAE (see Table 1)

between V_β expression and T-cell specificity[28] The TCR genes from encephalitogenic T-cell clone PJR-25 are being cloned, and should be useful probes allowing us to examine whether other encephalitogenic clones, which share the same epitope and class II restriction, express the same TCR genes

We have described here the first investigation of the fine specificity of T cell clones mediating an autoimmune disease Two distinct T-cell epitopes in MBP 1-16 have been characterized Examination of the response to MBP peptides, both *in vitro* and *in vivo*, has revealed that only a limited repertoire of N-terminal specific T lymphocytes, expressing a common class II restriction and peptide specificity, mediate induction of EAE These results may provide insights into the immunogenetic mechanisms of other autoimmune diseases In cases where only a limited repertoire of T lymphocytes mediate autoimmune pathogenesis, specific forms of therapy might be possible For example measures which selectively inactivate or eliminate only autoaggressive T cells bearing the N-terminal MBP specificity may be beneficial in the treatment of autoimmune encephalomyelitis

We thank Teri Montgomery for her valuable assistance and Professors Hans Acha-Orbea, Mark Davis, Robert Fritz, Garry Fathman, Phillipa Marrack, Hugh McDevitt, Patricia Nelson, Tim Vollmer and Matthew Waldor for helpful discussions I S is supported by NIH grants NS00571, AI22462 and NS18235, the Rosenthal Foundation, the Fausel Foundation and the Swim Foundation

Received 16 May, accepted 6 August 1986

1 Wisniewski, H M & Keith, A B *A. Neurol.* 1, 144-148 (1977)
2 Raine, C S, Snyder, D H, Valsamis, M P & Stone, S H *Lab Invest* **31,** 369-380 (1974)
3 Raine, C S in *Multiple Sclerosis—Pathology, Diagnosis and Management* (eds Hallpike, J, Adams, C W M & Tourtellote, W W) 413-460 (Chapman & Hall, London, 1983)
4 Lassman, H K & Wisniewski, H M *Archiv Neurol.* **36,** 490-497 (1979)
5 Zamvil, S *et al* *Nature* **317,** 355-358 (1985)
6 Zamvil, S *et al* *J exp Med.* **162,** 2107-2124 (1985)
7 Petinelli, C B & McFarlin, D E *J Immun.* **127,** 1420-1423 (1981)
8 Brostoff, S W & Mason, D W *J Immun.* **133,** 1938-1942 (1984)
9 Waldor, *et al* *Science* **227,** 415-417 (1985)
10 Petinelli, C B, Fritz, R B, Jen Chou, C H & McFarlin, D E *J Immun* **129,** 1209-1214 (1982)
11 Fritz, R. B, Jen Chou, C H & McFarlin, D E, *J Immun* **130,** 191-194 (1983)
12 Fritz, R B, Skeen, M. J, Jen Chou, C H, Garcia, J & Egorov, I K, *J Immun* **134,** 2328-2332 (1985)
13 Schwartz, R H, Fox, B S, Frago, E, Chen, C & Singh, B *J Immun* **135,** 2598-2608 (1985)
14 Haskins, K *et al.* *J exp Med* **157,** 1149-1169 (1983)
15 Kaye, J, Porcelli, S, Tite, J Jones, B & Janeway, C A Jr *J exp Med* **158,** 836-856 (1983)
16 Martenson, R E. *Prog clin biol. Res* **146,** 511-521 (1984)
17 Eylar, E H, Caccam J, Jackson, J J, Westall, F C. & Robertson, R. B *Science* **168,** 1220-1223 (1970)
18 McFarlin, D E, Blank, S E, Kibler, R F, McKneally, S & Shapira, R *Science* **179,** 478480 (1972)
19 Teitlebaum, D C, Webb, C, Arnon, R & Sela, M *Cell. Immun* **29,** 265-271 (1977)
20 Kibler, R. F *et al.* *J exp Med.* 146, 1323-1331 (1977)
21 Hashim, G A, Carvalho, E F & Shapiro, R. D *J Immun.* **121,** 665-670 (1978)
22 Arnon, R *Immun. Rev* **55,** 5-30 (1981)
23 Yanagi, Y *et al.* *Nature* **308,** 145-149 (1984)
24 Hedrick, S M, Cohen, D I, Nielsen, E A & Davis, M M *Nature* **308,** 149-153 (1984)
25 Chien, Y *et al.* *Nature* **312,** 31-35 (1984)
26 Yague, J *et al.* *Cell* **42,** 81-87 (1985)
27 Fink, P J, Matis, L. A, McEllingott, D L, Bookman, M & Hendrick, S M *Nature* **321,** 219-226 (1986)
28 Goverman, J *et al.* *Cell* **40,** 859-867 (1985)
29 Erickson, B W & Merrifield, R B in *The Proteins* Vol II (ed Neurath, H) 255-527 (Academic, New York, 1976)
30 Rothbard, J B, Fernandez, R. & Schoolnick, G K *J exp Med.* **160,** 208-221 (1984)
31 Kimoto, M & Fathman, C G *J exp Med.* **152,** 759-770 (1980)
32 Oi, V T, Jones, P P, Goding, J W & Herzenberg, L A *Clin. Topics Microbiol. Immun.* **81,** 115-129 (1978)
33 Ozato, K, Mayer, N & Sachs, D H *J Immun* **124,** 533-540 (1980)

Interaction of peptide antigens and class II major histocompatibility complex antigens

Jean-Gerrard Guillet, Ming-Zong Lai, Thomas J. Briner, John A. Smith* & Malcolm L. Gefter

Department of Biology, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA
* Departments of Molecular Biology and Pathology,
Massachusetts General Hospital and Department of Pathology,
Harvard Medical School, Boston, Massachusetts 02114, USA

T lymphocytes require a foreign antigen to be presented on a cell surface in association with a self-transplantation antigen before they can recognize it effectively. This phenomenon is known as major histocompatibility complex (MHC) restriction[1]. It is not clear how an incalculably large number of foreign proteins form unique complexes with a very limited number of MHC molecules. We studied the recognition properties of T cells specific for a peptide derived from bacteriophage λ cI protein. Analogues of this peptide, as well as peptides derived from other unrelated antigens which can be presented in the context of the same MHC molecule, can competitively inhibit activation of these T cells by the cI peptide. Furthermore, these unrelated antigens can stimulate cI-specific T cells if certain specific amino-acid residues are replaced. Here we suggest a model in which all antigens give rise to peptides that can bind to the same site on the MHC molecule. T-cell recognition of this site (which is presumed to be polymorphic) with or without antigen bound can explain self-selection in the thymus and MHC restriction.

We immunized BALB/c mice with the 102 amino-acid N-terminal domain of bacteriophage λ cI protein (cI) and produced cI-specific, class II-restricted T-cell hybridomas More than 75% of specific hybridomas could be stimulated by a peptide comprised of residues 12-26 (P12-26) in the context of 1-A^d Independently isolated T cells could be divided into at least three groups on the basis of their reactivity to truncated analogues of P12-26 (Fig 1) Following Rock and Benacerraf[2] and others[3], who showed competition for presentation between related antigens, we examined whether inactive truncated analogues could inhibit activation of T cells stimulated with the parent peptide or active truncated analogues In the case of the hybridoma 7B7 3, P12-24 can inhibit activation by P15-26, but not in the case of hybridoma 9H35 (see Fig 2*a*) Failure to obtain competition has also been noted in other systems[1,4], suggesting that competition depends on the T cell involved It may well be that the active peptide and the analogues compete for binding to the same site on I-A (ref 5) but the T cell enhances the avidity of the active peptide for I-A so that competition cannot be seen This hypothesis agrees with recent findings from several laboratories[6,7] To test the hypothesis that other antigens which can be presented in the context of 1-A^d inhibit cI peptide stimulation, we screened 14 overlapping peptides derived from

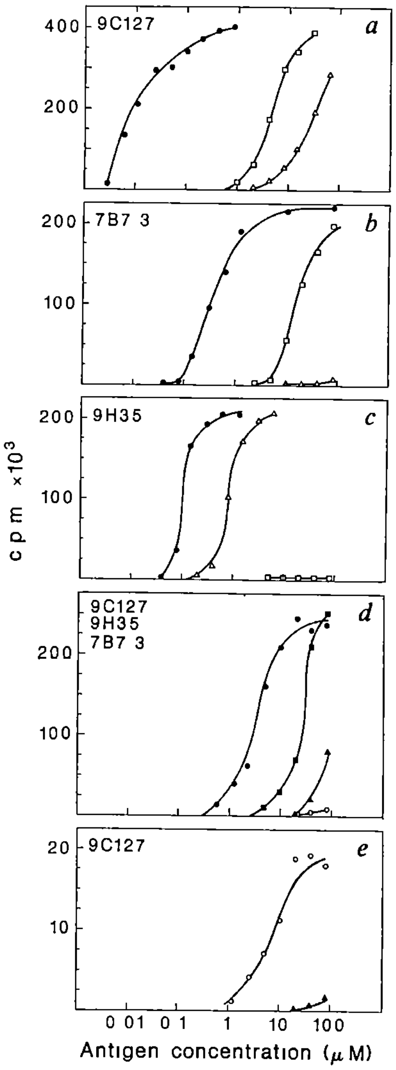


Fig. 1 *a–c*, Response of cI-specific T-cell hybridomas to various peptides derived from bacteriophage cI protein. Activity was measured with P12-26 (residues 12-26) (●), P15-26 (residues 15-26 (□) and P12-24 (residues 12-24) (△) as measured by IL-2 production. *a*, Hybridoma 9C127, *b*, hybridoma 7Bn,3, *c*, hybridoma 9H35. Values are the arithmetic mean of triplicate samples. *d*, *e* The stimulation of cI-specific T cells by analogues of peptides derived from *S aureus* nuclease and chicken ovalbumin. *d*, Hybridomas 9C127 (●), 9H35 (▲) and 7B7 3 (■) were incubated in the presence of [Lys 73, Ile 75] Nase I (see Fig 3) at the concentration indicated and activation was measured by IL-2 release as described below. The activity in the presence of Nase I (see Fig 3) for all the T-cell hybridomas is also shown (○). *e*, Hybridoma 9C127 was incubated in the presence of [Lys 331, Ile 333, Lys 336] Ova I (○) (see Fig 3) and Ova I (▲) (see Fig 3), as described above.

Methods. The T-cell hybridomas were derived from BALB/c mice immunized with bacteriophage λ repressor cI protein N-terminal fragment (residues 1-102), as described[4]. Those T cells reacting to a synthetic peptide (residues 12-26) were selected. The complete sequence of λ repressor cI protein is reported by Sauer and Anderegg[13]. All peptides used in this letter were made by the solid phase method of Merrifield[14] using an automated Applied Biosystems 430A peptide synthesizer. The peptides were desalted on a G-25 column (2.5×90 cm) equilibrated with 4 M glacial acetic acid and then lyophilized twice after resuspension in distilled water. 5×10^4 A20 antigen presenting cells were incubated in 200 µl in 96-well Costar plates with 5×10^4 hybridoma T cells at the indicated concentration of peptide. After 24 h of culture, supernatant (50 µl) was harvested and assayed for IL-2 concentration by the following incorporation of ^{3}H-thymidine into the IL-2 dependent CTL-L cell line (10^4 cells per well). Twenty-four hours after culture of T-cell supernatant in the presence of CTL-L cells, 1 µCi of ^{3}H-thymidine was added per well. Six hours later the cells were harvested by an automated cell harvester (Skatron Inc.) and thymidine incorporation was measured by scintillation counting.

the antigen *Staphylococcus aureus* nuclease (Nase)[8]. One peptide (residues 61-80) was a potent inhibitor of all three cI-specific T cells tested (Fig 2*b*). Note that the stimulating analogue showing the lowest apparent avidity was used for each T cell respectively (see Fig 2, legend) to allow observation of inhibition.

If the inhibition observed with the Nase peptide is biologically relevant, that is, if its ability to bind to a specific site on the I-A molecule is related to its ability to be presented, then Nase peptide P61-80 should also be the target peptide for Nase reactivity in the H-2^d mouse. This is indeed the case[8]. As a control, the Nase peptide P81-100 is the target peptide for activation in the context of I-E^k and does not act as an inhibitor in this system[8]. Examination of the sequence of residues of 61-80 revealed that residues 66-78 (Nase I) were remarkably similar to the cI peptide P12-24 (Fig 3). There are only two residues (positions 73 and 75) that differ significantly, all other changes are conservative. If Nase I binds to exactly the same site as P12-24 and if the structures of the bound peptides are the same, we reasoned that replacement of the residues at positions 73 and 75 of Nase I with the residues found in the equivalent positions on P12-24 might produce a peptide able to stimulate the cI-specific T cells. This is indeed the case (Fig 1*d*). Both peptides are equally active for the T cell 9C127, but the T cell 9H35 requires ~100-fold higher concentration of the Nase analogue. Surprisingly, the peptide [Lys 73, Ile 75] Nase I stimulates the T cell 7B7 3, whereas the cI derived peptide (the immunogen analogue) P12-24 does not. The unmodified Nase peptide does not stimulate any cI-specific T cells.

This suggested to us that all peptides restricted by I-A^d might bind to exactly the same site on I-A and act as inhibitors of cI-specific T cell activation. The peptide derived from ovalbumin, residues 324-336 (Ova I) is I-A^d restricted, and its behaviour has been well characterized[4,9]. As shown in Fig 2*c*, this peptide inhibits all cI-specific T cells tested. However, this peptide has little obvious similarity to P12-24 (Fig 3). We have shown independently[10] that P12-26 competes with Ova P323-339 for binding to the I-A molecule. We assume that Ova I is capable of forming a structure very similar to P12-24 either in solution or, more likely, once bound to the I-A molecule. If we assume that both peptides bind to the same site on I-A and that the structural requirements for T-cell activation are the same for all antigens, then we should be able to make Ova I activate cI-specific T cells by replacing a few critical residues on the Ova I peptide. Watts *et al.*[9] have shown that the histidine residue (boxed in Fig 3) is a critical 'epitope' for Ova-specific T-cell activation. Using this residue to align Ova I with P12-24, we made the peptide [Lys 331, Ile 333, Lys 336] Ova I (see Fig 3). This peptide can stimulate the cI-specific T-cell hybridoma 9C127 (Fig 1*e*), which has the highest affinity for 12-24. The concentration of the Ova analogue peptide required to give half-maximal stimulation of 9C127 is the same as P12-24. This probably reflects its intrinsic affinity for I-A (also as measured in the inhibition assay). The maximal level of stimulation for the 9C127 T cell is 15-20% of that seen with P12-24, this probably reflects a weak affinity of the T cell for the bound antigen. It is plausible that maximal T-cell stimulation requires

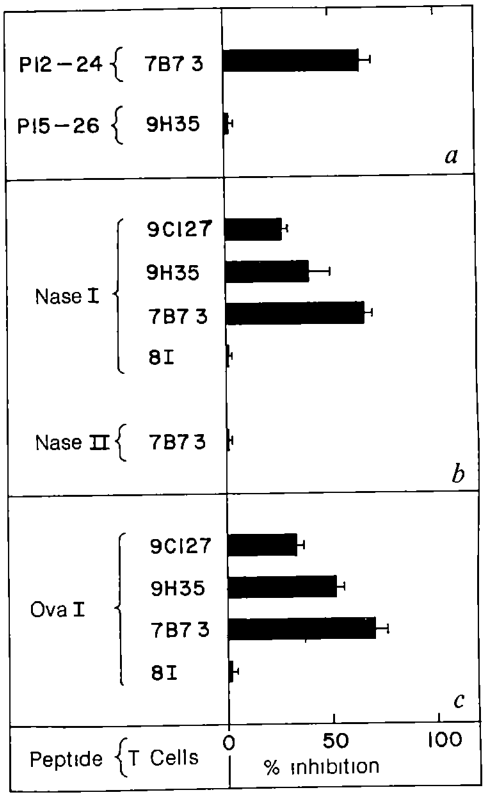


Fig. 2 Inhibition of T-cell activation by various peptides The extent of inhibition was analysed by comparing the amount of IL-2 production in the presence of a fixed concentration of inhibitor and varying the amount of stimulating peptide over a 50-fold range Inhibition was clearly competitive in that the extent of inhibition seen was greatest at high ratios of inhibitor to activator and least when the ratios were reversed For simplicity, we present here the inhibition data obtained at a concentration of stimulator which gives 50% of maximal stimulation, the inhibitor concentrations are given below *a*, The activity of T-cell hybridoma 7B7 3 was measured in the presence of P15-26 (20 μM) alone as described in the legend to Fig 1 or in the presence of both P15-26 (20 μM) and P12-24 (40 μM) The response of T-cell hybridoma 9H35 was measured in the presence of P12-24 (1 μM) alone or in combination with P15-26 (60 μM) The difference in activity seen when two peptides were included in assay cultures compared to the activity seen when stimulator peptide alone was given is presented as the per cent inhibition *b*, The activity of the T-cell hybridoma 9C127 was measured in the presence of P12-24 (20 μM) alone, as described in Fig 1 (legend) or in the presence of both P12-24 (20 μM) and a synthetic peptide (residues 61-80) derived from the protein *S aureus* nuclease (Nase I) (60 μM) The response of the hybridoma 9H35 was measured in the presence of P12-24 (1 μM) alone or in combination with Nase I (60 μM) The response of the hybridoma 7B7 3 was measured in the presence of P15-26 (20 μM) alone or in combination with Nase I (60 μM) or Nase II (residues 81-100) (100 μM) The activity of T-cell hybridoma 8I was measured in the presence of P15-26 (0 25 μM) alone or in combination with Nase I (60 μM) T-cell hybridoma 8I was derived from immune A/J mice and recognized P15-26 in the context of I-E^k Accordingly the cell line TA3 was used for antigen presentation *c*, The same conditions as *b*, except Ova I (25 μM) was used as an inhibitor

Fig. 3 The wild-type sequence of bacteriophage λ repressor protein residues 12-24 (P12-24), is shown Nase I represents residues 66-78 from *S aureus* nuclease and Ova I represents residues 324-336 from chicken ovalbumin The residues in boxes indicate the locations where modifications to the parent peptides were made All peptides and their analogues were synthesized and purified as described in Fig 1 The peptide [Lys 73, Ile 75] Nase I has the amino-acid sequence of Nase I, except that the residues at positions 73 and 75 were replaced by the residues present at the boxed positions of P12-24 The peptide [Lys 331, Ile 333, Lys 336] Ova I has the amino-acid sequence of Ova I, except that the residues at positions 331, 333 and 336 were replaced by the residues at the boxed positions of P12-24

P12-24	LEU -	GLU -	ASP -	ALA -	ARG -	ARG -	LEU -	LYS -	ALA -	ILE -	TYR -	GLU -	LYS
Nase I	VAL -	GLU -	ASN -	ALA -	LYS -	LYS -	ILE -	GLU -	VAL -	GLU -	PHE -	ASP -	LYS
Ova I	SER -	GLN -	ALA -	VAL -	HIS -	ALA -	ALA -	HIS -	ALA -	GLU -	ILE -	ASN -	GLU

more than the two identical amino-acid contacts provided by the analogue

These results suggest that the class II transplantation antigen is a simple receptor containing a single binding site for peptides able to form similar general structures once bound The function of the transplantation antigen would be to position the antigen for T-cell binding uniquely It has been postulated that all T-cell antigens are peptides capable of forming amphipathic α helices[11], it is not likely that the peptide P12-26 described here will do so However, it is attractive to consider that a structure like an α helix or β sheet is the common structure that immunogenic peptides form either in solution, or, more likely, when bound to the transplantation antigen Given recent results[12], it is also reasonable to propose that similar models can be constructed for class I-restricted antigen presentation The 'docking' of the T-cell receptor onto the transplantation antigen which allows it to interact with the antigen uniquely requires an intrinsic affinity of the T-cell receptor for the area of the protein surrounding the peptide binding site This area is the same for all antigens presented on a given transplantation antigen molecule, and is presumably the site learned as 'self' in the thymus It follows that the binding site and its local area is the polymorphic region of the I-A molecule which imparts unique qualities to both the peptide binding site itself (controlling which amino-acid sequences bind best to the site) and the self recognition site(s) We envisage that the transplantation antigen has one peptide binding site The properties of the site (α helical-binding) would be imposed by the general structure of the transplantation antigen (class I or class II)

This model provides a framework for the construction of synthetic immune agonists and antagonists Further physical and chemical data are required to test these proposals

This work is supported by grant AI13357 from NIH to M L G and by a grant from Hoechst Aktiengesellschaft (FRG) to J A S M Z L is supported by a postdoctoral fellowship, DRG-853, from the Damon Runyon-Walter Winchell Cancer Fund

Received 2 July, accepted 11 September 1986

1 Schwartz, R. H *A. Rev Immun* **3**, 237-261 (1985)
2 Rock, K. L & Benacerraf, B *J exp Med* **157**, 1618-1634 (1983)
3 Werdelin, O *J Immun.* **129**, 1883-1891 (1982)
4 Shimonkevitz, R, Colon, S, Kappler, J W, Marrack, P & Grey, H M *J Immun* **133**, 2067-2074 (1984)
5 Babbitt, B P, Allen, P M, Matsueda, G, Haber, E & Unanue, E R *Nature* **317**, 359-360 (1985)
6 Watts, T H, Gaub, H E & McConnell, H M *Nature* **320**, 179-181 (1986)
7 Ashwell, S D & Schwartz, R H *Nature* **320**, 176-179 (1986)
8 Finnegan, A *et al J exp Med* **164**, 899-910 (1986)
9 Watts, T H, Ganepy, J, Schoolnick, G K & McConnell, H M *Proc. natn. Acad Sci USA* **82**, 5480-5484 (1985)
10 Guillet, J-G *et al. Proc. natn. Acad Sci USA* (submitted)
11 Delisi, C & Berzofsky, J A *Proc natn Acad Sci USA* **82**, 7048-7052 (1985)
12 Townsend, A R M *et al. Cell* **44**, 959-968 (1986)
13 Sauer, R T & Anderegg, R *Biochemistry* **17**, 1092-1100 (1978)
14 Merrifield, R B, *J Am. chem Soc* **85**, 2149-2154 (1963)

Identification of a common class of high affinity receptors for both types of porcine interleukin-1 on connective tissue cells

T. A. Bird & J. Saklatvala

Strangeways Research Laboratory, Worts Causeway,
Cambridge CB1 4RN, UK

Interleukin-1 (IL-1) is the name given to the polypeptides produced by activated mononuclear phagocytes which were originally defined as lymphocyte activating factors (LAF)[1]. Administration of IL-1 *in vivo* causes fever[2] and synthesis of acute phase proteins[3]. *In vitro* they have been shown to cause cartilage and bone resorption[4-7], and to stimulate fibroblasts and chondrocytes to make prostaglandins and latent collagenase[4,8]. IL-1 has therefore been proposed to be an important inflammatory mediator[9] and may be involved in the destruction of cartilage and bone that is a feature of rheumatoid arthritis and other inflammatory diseases of joints. We therefore looked for IL-1 receptors on connective tissue cells which might be targets for therapeutic intervention. Here we report the iodination, to high specific activity and with retention of full biological potency, of the two types of natural porcine IL-1. These ligands have been used to demonstrate high affinity dissociation constant ($\sim 10^{-10}$ M) specific binding sites on pig chondrocytes and synovial fibroblasts, human dermal fibroblasts and murine osteoblasts (3,000–5,000 sites per cell). Most interestingly, the two different Il-1 proteins show a similar affinity for a common class of receptors.

Two types of IL-1 have been isolated from culture supernatants of porcine mononuclear leucocytes stimulated with Concanavalin A. These are immunologically different proteins that have identical relative molecular masses (M_rs) of 21,000 (21K) on SDS polyacrylamide gel electrophoresis (PAGE) and have been characterized elsewhere[5]. One is acidic, having an isoelectric point (pI) of 5 (IL-1/5) and the other basic (pI of 8). These proteins are probably the porcine counterparts of the human IL-1α (which has a pI of 5) and the human IL-1β (which has a pI of 7 (ref. 10)) since an antiserum to porcine IL-1/5 cross reacts with human IL-1α and the N-terminal amino-acid sequence of porcine IL-1/8 shows 35% homology with that of human IL-1β (J.S., unpublished observations).

Attempts to label the different types of porcine IL-1 by direct oxidative methods proved unsuccessful, resulting in loss of biological activity. However, we were able to produce ^{125}I-labelled preparations of IL-1 which were fully active in a cartilage-resorption assay (Fig. 1*b*) using di-^{125}I-Bolton and Hunter reagent. The labelled proteins migrated as single bands upon SDS-PAGE, in positions expected for the native 21K proteins (Fig. 1*a*). The specific activity in each case was 470–580 Ci mmol^{-1}. ^{125}I-IL-1 could be stored at −20 °C for at least 6 weeks without loss of binding affinity. We initially studied the interaction of ^{125}I-IL-1/8 and ^{125}I-IL-1/5 with monolayers of pig synovial fibroblasts. Figure 2*a* shows the time course for specific (displaceable) binding of ^{125}I-IL-1/8 (2.4×10^{-10} M) to pig synovial fibroblasts at various temperatures. The rate of binding increased with increasing temperature, equilibrium was only attained after 6 h of incubation at 19 °C. To examine whether some ^{125}I-IL-1/8 was internalized at 37 °C parallel cultures were briefly exposed to an acidic glycine buffer[12] to remove surface-bound ligand. 85% of the radioactivity could be removed from monolayers incubated at 4 °C or 19 °C in this way but, as shown in Fig. 2*a*, a large proportion of the radioactivity bound at 37 °C rapidly became resistant to this treatment and was presumably internalized. The amount of acid-labile radioactivity bound at 37 °C reached a steady state within one hour, the fact that a

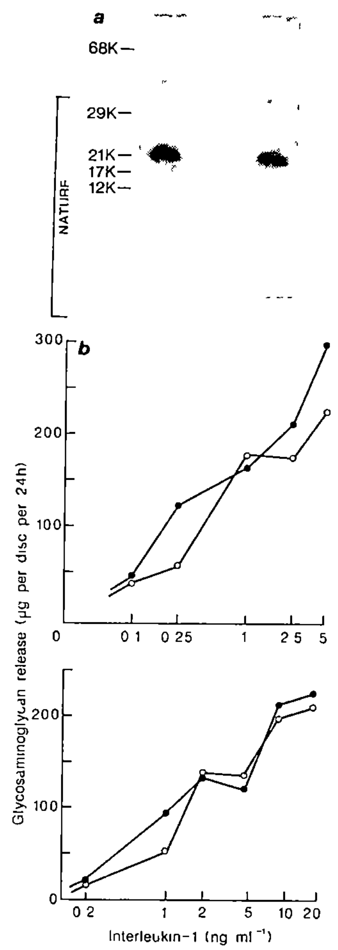


Fig. 1 *a*, Autoradiographs of ^{125}I-IL-1/8 (lane 1) and ^{125}I-IL-1/5 (lane 2) electrophoresed on 12.5% SDS-polyacrylamide slab gels in an ammediol buffer system[16] with reduction. The M_rs of Coomassie brilliant blue-stained standard proteins run in parallel are indicated in kilodaltons, *b*, comparison of ^{125}I-IL-1/5 (●, upper panel) and ^{125}I-IL-1/8 (●, lower panel) with the corresponding native interleukin 1 (○) in a cartilage resorption assay. Vertical axis, glycosaminoglycan release (µg per disc per 24 h), horizontal axis, concentration of IL-1 (ng ml^{-1}).

Methods. Purified IL-1s (2–4 µg estimated from A_{280}, assuming absorbance of a 1 mg ml^{-1} solution in a 1 cm lightpath to be unity) were iodinated by the method of Bolton and Hunter[17]. The interleukins, freeze-dried from aqueous solution, were dissolved in 20 µl 0.1 M borate buffer (pH 8.5) and added to 1 mCi *N*-succinimidyl 3-(4 hydroxy-3,5-[^{125}I] di-iodophenyl) propionate freshly dried from a solution in benzene. After agitation of the mixture on ice for 30 min the reaction was quenched by addition of 50 µl 1 M glycine in borate buffer. Iodinated proteins were separated from byproducts by gel filtration on Sephadex G-25 which was equilibrated and eluted with 0.05 M sodium phosphate buffer (pH 7.5) containing 0.25% (w/v) gelatin. Because ^{125}I-glycine and free ^{125}I-Bolton and Hunter reagent were quantitatively eluted from the column, residual radioactivity in the gel slurry was counted to estimate retained ^{125}I-IL-1. Recovery of labelled protein was typically 75%. After correcting for recovery, aliquots of labelled IL-1 were tested against unlabelled material for their ability to stimulate cartilage resorption[4]. Disks (2 mm) of bovine nasal cartilage were maintained in 96-well microtitre plates in 150 µl of Dulbecco's modified Eagles Medium (DMEM) containing 5% (v/v) bovine serum and test samples for 24 h at 37 °C in CO_2/air (1:19). Chondroitin sulphate released into the medium was measured using the metachromic dye dimethyl methylene blue. Each point represents the mean of four disks. Release of chondroitin sulphate in the absence of interleukin 1 was less than 30 µg per disk. Note that the different dose/response relationships in Fig. 1*b* arise from the fact that different batches of cartilage were used for each assay. When the two types of IL-1 are tested on the same batch of cartilage, they have the same specific bioactivity[5].

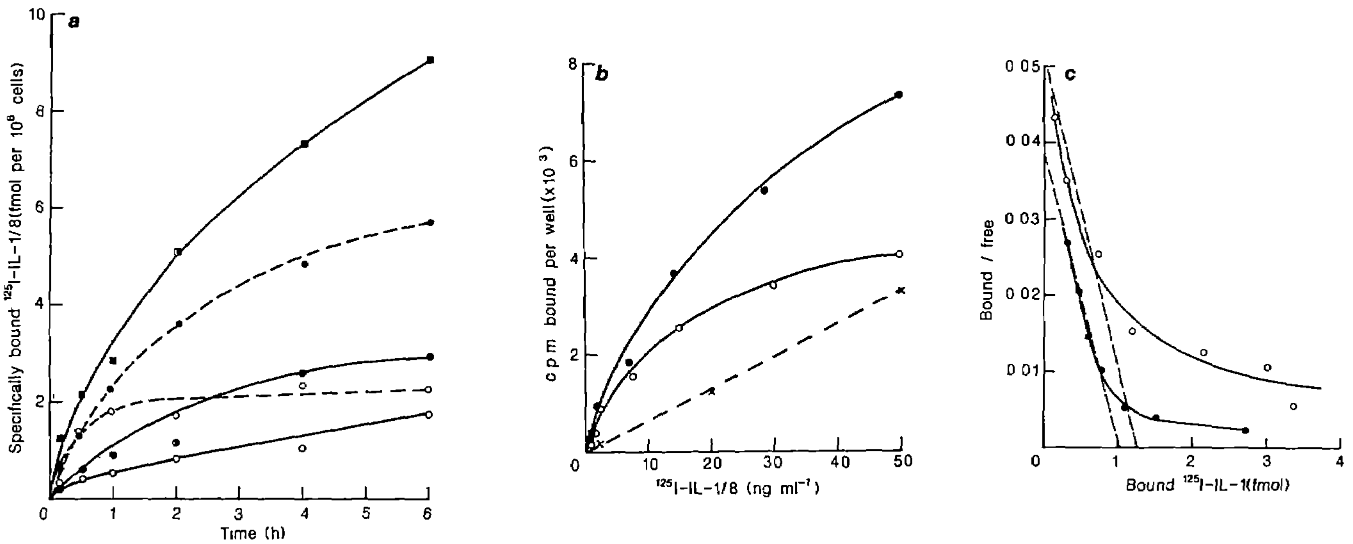


Fig 2 *a*, Time course of ^{125}I-IL-1/8 binding to porcine synovial fibroblasts (PSF) at various temperatures, and internalization of ligand at 37 °C ■—■, binding at 37 °C, ●—●, binding at 19 °C, ○—○, binding at 4 °C, ○---○, acid-labile radioactivity, ●---●, residual, SDS soluble radioactivity, *b*, saturable binding of ^{125}I-IL-1/8 to PSF incubated for 4 h at 19 °C, ●, total binding, ○, specific binding, ×, nonspecific binding *c*, Scatchard analysis of ^{125}I-IL-1/5 (●) and ^{125}I-IL-1/8 (○) to PSF at 19 °C

Methods. *a*, *Monolayers of PSF, isolated from metacarpophalangeal joints of freshly-killed young pigs by enzymatic dissagregation*[18] were maintained in DMEM plus 10% (v/v) heat-inactivated fetal calf serum until required Newly confluent monolayers ($1\ 2\times10^5$ cells) in 4 cm^2 plastic wells were washed twice with Hank's balanced salts solution supplemented with 0 1% BSA and 20 mM HEPES *p*H 7 4 (binding buffer), then incubated on a rotatory platform (70 cycles min^{-1}) *for the indicated times at* 4 °C, 19 °C or 37 °C with 400 µl per well binding buffer containing $2\ 4\times10^{-10}$ M ^{125}I-IL-1/8 Monolayers were then rinsed with 3 changes of ice cold binding buffer (1 ml), solubilized with 0 5 ml 1% SDS and counted in a gamma counter In some cases, monolayers incubated with ^{125}I-IL-1 at 37 °C were rinsed with binding buffer and kept on ice for 15 min in 1 ml of 50 mm glycine/150 mM NaCl *pH* 3 0 *before solubilization in* 1% SDS The acid-labile radioactivity and residual, SDS soluble radioactivity were counted separately Each point is the mean of three determinations from which nonspecific binding (measured in the presence of 5×10^{-8} M partially purified IL-1) has been subtracted Conditions in *b* and *c* were the same with varying ^{125}I-IL-1/8 *concentrations at* 19 °C *for* 4 h Each *point is calculated from the average values of four determinations Straight lines were fitted,* where appropriate, by the method of least squares

pool of surface-bound ligand was maintained despite the appearance of radioactivity in an intracellular compartment suggests receptor recycling such as that described for the LDL receptor[13] For subsequent studies, cells were incubated with labelled IL-1 at 19 °C for 4 hours Figure 2*b* shows that binding of ^{125}I-IL-1/8 to pig synovial fibroblasts at 19 °C was saturable, with half maximal binding occurring at 4×10^{-10} M (9 ng ml^{-1}), nonspecific binding was 10–30% of total binding Similar data were obtained using ^{125}I-IL-1/5 Scatchard analysis of binding to pig synovial fibroblasts (Fig 2*c*) gave curvilinear plots resolving into a high affinity component and, usually, a much lower affinity component Straight lines could be fitted to points lying on the first part of the plots to yield, by extrapolation, capacities of approximately 5,400 and 4,500 sites per cell for ^{125}I-IL-1/8 and ^{125}I-IL-1/5 respectively, with equilibrium dissociation constants (K_D of $1\ 5\times10^{-10}$ and $1\ 7\times10^{-10}$ M) High affinity binding was reproducible over a number of experiments using different batches of cells and different preparations of ^{125}I-IL-1, but the contribution of low affinity binding was less consistent Given the sensitivity of Scatchard analysis to various sources of error at high levels of bound ligand[14], this component may be artefactual and result from imprecise measurement of non-specific binding Similar experiments were performed using human dermal fibroblasts and porcine articular chondrocytes In each case linear Scatchard plots were obtained, human fibroblasts bind 4,800 sites per cell with a K_D of $6\ 8\times10^{-11}$ M, porcine chondrocytes bind 7,000 sites per cell with a K_D of $2\ 5\times10^{-11}$ M Reversibility of binding was assessed after preincubation of pig synovial fibroblasts with 10 ng per ml ^{125}I-IL-1/8 (with or without excess unlabelled IL-1) The cells were washed and incubated at 19 °C in 2 5 ml per well of binding buffer When the preincubation was performed at 19 °C, 33% of specifically bound ligand was released over 6 hours The dissociation curve followed a first order model with $K_{-1}=2\times10^5$ s^{-1} and $t=$ 637 min The dissociable radioactivity was TCA-precipitable (not shown) 55% of the radioactivity bound at 37 °C became dissociated over 6 hours of which 40% was TCA-precipitable, indicating that some degradation had occurred

We wished to test whether both types of ^{125}I-IL-1 interacted with the same receptor sites, Fig 3*a* shows the results of experi-

Table 1 *Single-point equilibrium analysis of binding of* ^{125}I-IL-1/5 *and* ^{125}I-IL-1/8 *to various cells*

Cell type	Derivation	Species	Molecules bound per cell IL-1/8	IL-1/5
HuGi	Gingival fibroblast	Human	ND	715±45
Flow 7000	Foreskin fibroblast	Human	3,709±287	3,429±120
PSF	Synovial fibroblast	Pig	3,709±223	2,546±104
PCC	Chondrocyte	Pig	4,050±270	ND
PMO	Osteoblast	Mouse	2,605±800	3,248±122
3T3	Fibroblast	BALB/c Mouse	ND	2,752±146
HeLa	Cervical carcinoma	Human	ND	<100
HT1080	Fibrosarcoma	Human	<100	<100
MG63	Osteogenic sarcoma	Human	929±195	909±56
UMR 106	Osteogenic sarcoma	Rat	543±59	ND
M-3	Melanoma Cloudman	Mouse	<100	<100

Chondrocytes were isolated by sequential enzyme digestion[19] of pig metacarpal head cartilage and used in primary culture Pig synovial fibroblasts, prepared as described in Fig 1 (legend), were used between passages three and nine Human gingival fibroblasts, a gift from S J Atkinson, were subcultured through six passages before use Human foreskin fibroblasts (Flow 7000) were used between passages 17 and 21 Neonatal mouse calvarial osteoblasts, prepared by a sequential digestion procedure[20] were obtained *from Dr J Heath and used in primary culture The UMR-106 cell line* was obtained from Professor T Chambers, all other established lines were from Flow Laboratories All cells were maintained in DMEM supplemented with 10% (v/v) heat-inactivated FCS, and were subcultured, where appropriate, using trypsin-EDTA Confluent monolayers, in 4 cm^2 wells, were incubated for 4 h at 19 °C with ^{125}I-IL-1/8 ($3\ 57\times10^{-10}$ M) or ^{125}I-IL-1/5 ($2\ 38\times10^{-10}$ M) Specifically bound radioactivity was measured as described for Fig 2 Cells in additional wells were counted in a haemocytometer after trypsin-EDTA treatment Results are the mean ± s e m of four determinations, ND, not determined

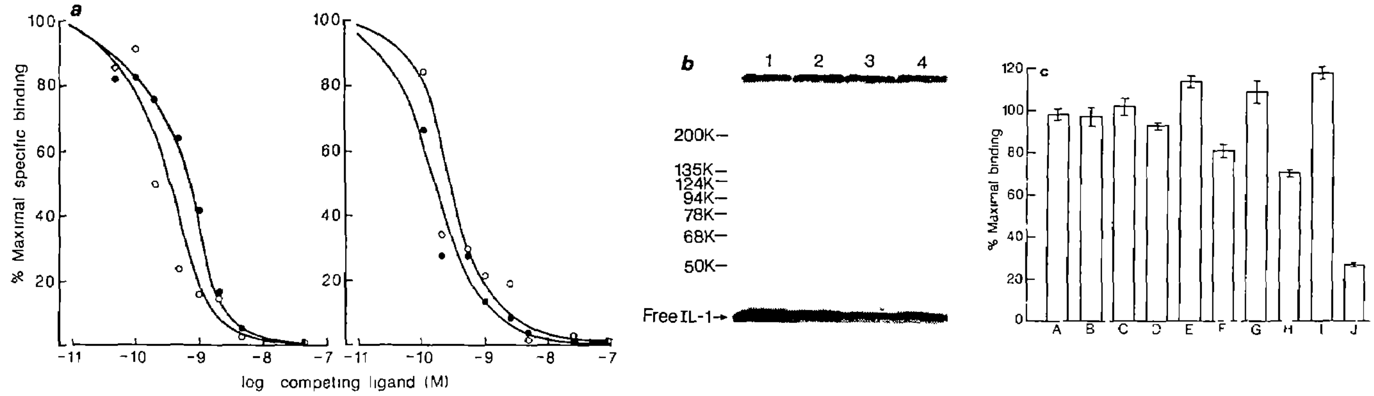


Fig 3 *a*, Competition of IL-1/5 (○) and IL-1/8 (●) with left panel, ^{125}I-IL-1/5, right panel, ^{125}I-IL-1/8 for binding to porcine synovial fibroblasts Vertical axis, per cent maximal specific binding, horizontal axis, $\log_{10}$ concentration of competing ligand *b*, Covalent crosslinking of ^{125}I-IL-1/8 to intact monolayers *c*, Specificity of the IL-1 receptor, the percentage of maximum ^{125}I-IL-1 binding observed in the presence of various other unlabelled polypeptide hormones and growth factors is shown

Methods Experimental conditions in *a* were as for Fig 2*b* with the addition of the indicated concentrations of unlabelled IL-1, labelled IL-1 was added at $1\ 2\times10^{-10}$ M Results are expressed as specific binding relative to binding in the absence of competing ligand and are the means of triplicate determinations *b*, Monolayers of 3T3 fibroblasts (clone A41, Flow Laboratories) in 25 cm^2 flasks were incubated for 4 h at 19 °C with 5 ml binding buffer supplemented with $9\ 5\times10^{-10}$ M ^{125}I-IL-1/8 and unlabelled IL-1/5 or IL-1/8 as indicated below Unbound radioactivity was removed by washing the monolayers with three changes of ice-cold protein-free binding buffer The flasks each received 5 ml of this buffer followed by 200 μl ethylene glycol bis (succinimidyl succinate) (Pierce) freshly dissolved in dimethyl sulphoxide, final crosslinker concentration was 0 5 mM After standing the flasks on ice for 45 min, the crosslinker was removed and replaced with 5 ml of a quenching solution (10 mM TrisHCl *p*H 7 5, 1 mM EDTA) The cells were scraped into this buffer, pelleted by centrifugation and dissolved in 300 μl of 1% SDS containing the protease inhibitors PMSF (10 mM) EDTA (2 mM) and *N*-ethylmaleimide (10 mM) Samples containing equal amounts of radioactivity were adjusted to 125 mM Tris HCl *p*H 6 8, 2% SDS, 5% 2-mercaptoethanol, boiled for 5 min and electrophoresed on SDS-polyacrylamide (5% stacking gel, 7 5% resolving gel)[21] The gel was dried under vacuum and exposed to X-ray film for 7 days at −80 °C Track, 1, no competing ligand, track 2, porcine IL-1/8 (1 μg ml^{-1}), track 3, porcine IL-1/5 (1 μg ml^{-1}), track 4, recombinant murine IL-1 (1 μg ml^{-1}) The positions and M_rs (in kilodaltons) of α2-macroglobulin, chick skin collagen, phosphorylase a, transferrin, serum albumin and goat IgG are given *c*, Porcine synovial fibroblasts were incubated with $2\ 4\times10^{-10}$ M ^{125}I-IL-1/5 and various unlabelled polypeptide hormones and growth factors A, synthetic human angiotensin II (2 μg ml^{-1}), B, human vasopressin (2 μg ml^{-1}), C, bovine parathyroid hormone (2 μg ml^{-1}), D, salmon calcitonin (2 μg ml^{-1}), E, mouse submaxillary gland epidermal growth factor (2 μg ml^{-1}), F, human platelet-derived growth factor (1 μg ml^{-1}), G, bovine insulin (3 μg ml^{-1}), H, mouse submaxillary gland 2 5 S nerve growth factor (90% pure, containing some renin-like activity 2 μg ml^{-1}), I, recombinant human tumour necrosis factor (2 μg ml^{-1}), J, porcine Il-1/5 (1 μg ml^{-1}) Values are expressed as the means ±s e m of radioactivity bound in triplicate determinations relative to an assay without added competitor Samples A, B, and G were from Sigma, E and F were from Bethesda Research Laboratories, C and D were obtained from Dr J K Health, Cambridge, H was a gift from Dr R Lindsay, Sandoz Institute for Medical Research, London, and sample I was provided by Dr M Palladino of Genentech, San Fransisco

ments in which unlabelled IL-1 (both types) was used to displace labelled IL-1 ($1\ 2\times10^{-10}$ M) from pig synovial fibroblasts It is clear that the two different molecules bind to a common class of receptor sites with an approximately equal affinity Very similar results were obtained using human dermal fibroblasts and porcine articular chondrocytes (not shown) Additional evidence that both types of IL-1 interact with a common site on fibroblasts is provided by crosslinking studies on pig synovial fibroblasts and murine 3T3 fibroblasts Similar results are obtained with both cell types but because the efficiency of crosslinking on 3T3 is greater, only these results are shown (Fig 3*b*) When ^{125}I-IL-1/8 was bound to fibroblasts in the presence of a homobifunctional crosslinker, faint high molecular weight complexes (95K and 190K) could be seen after SDS-PAGE and autoradiography The patterns were similar when samples were analysed under reducing and non-reducing conditions Inclusion in the incubation medium of a high concentration of either unlabelled porcine IL-1 or murine recombinant IL-1 prevented the formation of these complexes indicating that they are involved in the specific binding of IL-1 of either type

We examined the specificity of the IL-1 receptor further by including a series of polypeptide growth factors and hormones in the binding assay (Fig 3*c*) Of the ten factors tested, only IL-1 was able to compete effectively for binding of ^{125}IL-IL-1/5 to pig synovial fibroblasts We have screened other cells for IL-1 binding (see Table 1), receptor expression varies from a high level in chondrocytes and osteoblasts which respond strongly to IL-1 by producing collagenase and prostaglandin (unpublished results), to cells without measurable binding (HeLa and HT1080 which do not produce PGE$_2$ in response to IL-1 and may be relatively insensitive Our findings for fibroblasts are quantitatively similar to those recently reported[15] for binding of human ^{125}I-IL-1 (*p*I7) which was 95% inactivated by Bolton-Hunter labelling fibroblasts to human and mouse fibroblasts, and they substantiate an initial report[22] of cross competition for binding of recombinant human IL-1 on murine T cells The present report indicates the suitability of the different types of porcine IL-1 as probes for receptors, as they can be radiolabelled preserving full activity, and provides the first evidence that the two forms show similar affinity for a common receptor on relevant target cells

We thank Dr M Palladino for the sample of recombinant TNF, Dr P Lomedico for the gift of recombinant murine IL-1 and the Arthritis and Rheumatism Council for their generous support

Received 29 January, accepted 10 September 1986

1 Mizel, S B , *Immun. Rev* **63,** 51–72 (1982)
2 Murphy, P A , Simon, P L R & Willoughby, W F *J Immun.* **124,** 2498–2501 (1980)
3 Sztein, M B *et al. Cell Immun.* **63,** 164–176 (1981)
4 Saklatvala, J , Pilsworth, L M C , Sarsfield, S J , Gavrilovic, J & Heath, J K *Biochem. J* **224,** 461–466 (1984)
5 Saklatvala, J , Sarsfield, S J & Townsend, V *J exp Med* **162,** 1208–1222 (1985)
6 Heath, J K , Saklatvala, J , Meikle, M C , Atkinson, S J & Reynolds, J J , *Calc. Tissue Int* **37,** 95–97 (1985)
7 Gowen, M , Wood, D D , Ihrie, E J , McGuire, M K B & Russell, R G G *Nature* **306,** 378–380 (1983)
8 Mizel, S B , Dayer, J -M , Krane, S M & Mergenhagen, S E , *Proc. natn. Acad Sci USA.* **78,** 2474–2477 (1981)
9 Oppenheim, J J & Gery, I *Immun Today* **3,** 113–119 (1982)
10 March, C J *et al Nature* **315,** 641–647 (1985)

11 Saklatvala, J , Sarsfield, S J & Wood, D D *Br J Rheumat* **24**, (Suppl) 68-71 (1985)
12 Ascoli, M , *J biol Chem* **257**, 13306-13311 (1982)
13 Goldstein, J L & Brown, M S *J cell Sci Suppl* **3**, 131-137 (1985)
14 Mendel, C M & Mendel, D B , *Biochem J* **228**, 269-272 (1985)
15 Dower, S K , *et al J exp Med* **162**, 501-515 (1985)
16 Wyckoff, M , Rodbard, D & Chrambach, A *Analyt Biochem* **78**, 459-482 (1977)
17 Bolton, A E & Hunter, W M *Biochem J* **133**, 529-539 (1973)
18 Pilsworth, L M C & Saklatvala, J *Biochem J* **216**, 481-489 (1983)
19 Schwartz, E R. in *Skeletal Research* (eds Simmons, D J & Kunin, A S) 61-80 (Academic, New York, 1979)
20 Heath, J K , Atkinson, S J , Meikle, M C & Reynolds, J J *Biochim biophys Acta* **802**, 151-154 (1984)
21 Laemmli, U K *Nature* **227**, 680-685 (1970)
22 Dower, S K , Call, S M , Gillis, S & Urdal, D L *Proc natn Acad Sci U S A* **83**, 1060-1064 (1986)

The cell surface receptors for interleukin-1α and interleukin-1β are identical

Steven K. Dower, Shirley R. Kronheim, Thomas P. Hopp, Michael Cantrell, Michael Deeley, Steven Gillis, Christopher S. Henney & David L. Urdal

Immunex Corporation, 51 University Street, Seattle,
Washington 98101, USA

Interleukin-1 (IL-1) is a factor that can induce proliferation of murine T lymphocytes[1,2] and can elicit a variety of other biological responses. These include bone resorption[3], fibroblast proliferation[4], acute phase protein release from hepatocytes[5], cartilage breakdown[6] and fever[7]. This spectrum of activities is consistent with a role for IL-1 as a mediator of inflammation. Recently, sequence data have shown that there are at least two members of the IL-1 family; these distantly related proteins have been termed IL-1α and IL-1β[8–11]. We have found previously that both murine T cells and fibroblasts possess a relative molecular mass (M_r) ~80,000 (80K) plasma membrane receptor for human IL-1β[12,13]. We show here that the receptor for IL-1α on both murine and human cells is identical to that for IL-1β. This result raises the issue of what separation, if any, there might be between the biological activities of IL-1α and IL-1β.

Both IL-1α and IL-1β are encoded by complementary DNAs for proteins of ~30K[8–11], whereas the active forms found in macrophage culture supernatants are ~17K[3,14–16] The start positions within the precursors for murine IL-1 and human IL-1β (nIL-1β) have been established by N-terminal protein sequence analysis[8 10] Recombinant human IL-1α (rIL-1α) and IL-1β (rIL-1β) were produced in *Escherichia coli* from truncated cDNA clones encoding polypeptides identical to the secreted forms, and subsequently purified to homogeneity (S R K, unpublished) nIL-1β was purified from culture supernatants and labelled with di-iodo (^{125}I) Bolton-Hunter reagent as described previously[12,15] rIL-1β was labelled with Bolton-Hunter reagent, and rIL-1α was labelled with sodium (^{125}I) iodide using a modified chloramine-T method[17] For all three lymphokine preparations, radioiodination yielded specific activities in the range 1×10^{15} to 5×10^{15} c p m mmol^{-1} (0 4-2 atoms I per molecule protein) and SDS-polyacrylamide gel electrophoresis (PAGE) revealed a single labelled polypeptide of 17 5K (Fig 1), consistent with previously reported[3,10,14–16] values for IL-1

All these labelled preparations were >98% TCA precipitable, thus the ^{125}I was covalently bound to protein Comparison of biological activities on murine T lymphoma cells[18] between labelled and unlabelled IL-1 showed that rIL-1β, like nIL-1β[12], suffered some loss of activity on labelling, but rIL-1α retained 100% biological activity (1×10^{12} U mg^{-1}) after labelling and >95% could be bound by IL-1 receptor bearing cells

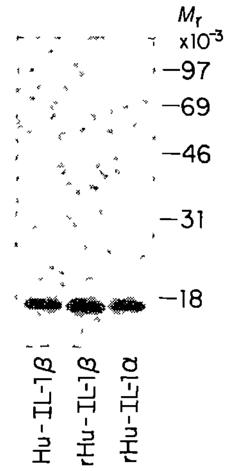


Fig 1 SDS-PAGE analysis of radiolabelled IL-1 preparations ^{125}I-IL-1 preparations were boiled for three minutes in sample buffer and analysed on a 10-20% polyacrylamide gradient slab gel as described elsewhere[12] Track 1, human IL-1β, track 2, recombinant human IL-1β, track 3, recombinant human IL-1α
Methods. IL-1 preparations to be used for radiolabelling were checked for purity and approximate protein concentration by SDS-PAGE and silver staining, the exact concentration was later determined by amino-acid analysis on an LKB 4150 amino-acid analyser and comparison with the predicted composition from the cDNA sequence[8] nIL-1β was labelled using di-iodo (^{125}I) Bolton-Hunter reagent (NEN) as described previously[12] Briefly 1 μg (0 06 nmoles) of nIL-1β in 50 μl of borate (0 02 M) buffered saline (0 15 M) was added to 10 μl of sodium borate (0 1 M) buffered saline (0 75 M) *p*H 8 5, the IL-1 was reacted with 1 mCi Bolton-Hunter reagent according to the manufacturer's instructions for 12 h at 8 °C Subsequently 30 μl of 2% gelatin and 5 μl of 1 M glycine ethyl ester were added and the protein separated from unreacted Bolton-Hunter reagent on a 1 ml bed volume Biogel P6 (Biorad) column Routinely ~5% of the Bolton-Hunter reagent was incorporated into protein Recombinant IL-1β was labelled with di-iodo Bolton-Hunter reagent by essentially the same protocol, except that 10 μg (0 57 nmol) of protein in 10 μl of phosphate (0 015 M) buffered saline (0 15 M) *p*H 7 2 were mixed with 10 μl of sodium borate (0 1 M) buffered saline (0 15 M) *p*H 8 5 and reacted with 1 mCi (0 23 nmoles) of Bolton-Hunter reagent Routinely 50-60% incorporation of label was observed Recombinant IL-1α was labelled with Na(^{125}I) using chloramine-T Briefly 10 μg of rIL-1α (0 57 nmol) in 10 μl of phosphate buffered saline were added to 2 5 mCi (1 0 nmol) of sodium iodide in 25 μl of 0 05 M sodium phosphate *p*H 7 0, reaction was initiated by addition of 30 μl of 1 4×10^{-4} M chloramine-T (4 2 nmol, Sigma) After 30 min on ice the reaction mixture was fractionated by gel filtration as described above Routinely 40-50% of ^{125}I was incorporated into protein Precipitation of the ^{125}I-IL-1 preparations in 1% gelatin carrier using 10% trichloroacetic acid, or rechromatography on gel filtration columns both showed that >98% of the ^{125}I label was protein bound after P6 purification

Table I summarizes the binding of ^{125}I-nIL-1β, ^{125}I-rIL-1β and ^{125}I-rIL-1α to LBRM-33-1A5 murine T lymphoma cells, human dermal fibroblasts, BALB-c/3T3 cells and ARH77, a human B lymphoblastoid cell line The number of sites on the LBRM-33-1A5 cells (Table 1) was at the high end of the range for 25 experiments, in which we observed considerable fluctuation in the number of receptors detected (290-3,430 sites per cell) The data in Table 1 show that all three types of ^{125}I-IL-1 detected a single class of high affinity sites on the cells Similar experiments with primary human dermal fibroblasts and murine 3T3 fibroblasts also showed that both IL-1α and IL-1β detected a single class of cell surface binding sites, the numbers of receptors detected by the two ligands were the same within experimental error (Table 1) In ARH77 cells, nIL-1β and rIL-1β bound to similar numbers of high affinity binding sites with identical affinities (Table 1) We could not obtain interpretable

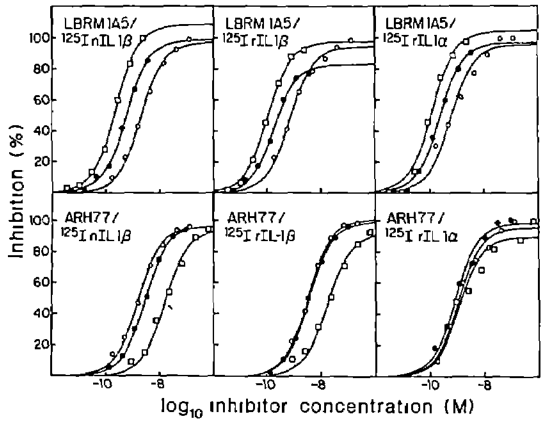


Fig. 2 Cross competition between rIL-1α, rIL-1β and nIL-1β on LBRM-33-1A5 and ARH77 cells Unlabelled competitor ●, nIL-1β, ○, rIL-1β, □, rIL-1α LBRM-33-1A5 cells ($3\ 1\times10^{7}$ per ml) or ARH77 cells ($2\ 1\times10^{7}$ per ml) were incubated with a fixed concentration of each type of ^{125}I-IL-1 and various concentrations of unlabelled IL-1 (as indicated in the figure) in 150 µl of RPMI 1640 with 1% BSA, 0 1% sodium azide and 20 mM HEPES *p*H 7 4 for 2 h at 8 °C on a rocker platform in round bottomed 96 well microtitre plates Concentrations of ^{125}I-IL-1 used were with LBRM-33-1A5 cells—nIL-1β, $3\ 9\times10^{-10}$ M, rIL-1β, $1\ 9\times10^{-9}$ M and rIL-1α, $9\ 2\times10^{-10}$ M, with ARH77 cells—nIL-1β, 9×10^{-10} M, rIL-1β, $1\ 9\times10^{-9}$ M and rIL-1α, $4\ 6\times10^{-9}$ M At the end of the assay, cells and bound radiolabel were separated from unbound label by centrifuging 60 µl aliquots of the reaction mixtures through 200 µl of phthalate oils in 400 µl polyethylene centrifuge tubes and excising the tips with a razor blade[17] Maximal binding was measured with the ^{125}I ligands alone, background binding was measured with ^{125}I-IL-1 and 6×10^{-7} M unlabelled homologous IL-1 All data have been corrected for nonspecific binding by subtraction The cells were greater than 90% viable by trypan blue exclusion at the end of the assay The continuous curves passing through the data were calculated from the best fit parameter values in Table 1 using a simple competitive inhibition model[19]

Table 1 Binding parameters for the interaction of rIL-1α, rIL-1β and nIL-1β with murine and human cells

Cell	Ligand	K_A (M^{-1})×10^{-9}	K_I (M^{-1})×10^{-9}	Receptors per cell (×10^{3})
LBRM-33-1A5	nIL-1β	1 4±0 1	4 0±1 8	2 8±0 1
LBRM-33-1A5	rIL-1β	0 3±0 02	1 8±0 3	3 2±0 2
LBRM-33-1A5	rIL-1α	4 6±0 7	11 3±2 1	3 43±0 02
ARH77	nIL-1β	0 44±0 06	0 7±0 4	11 3±0 3
ARH77	rIL-1β	0 5±0 1	1 5±0 6	11±1
ARH77	rIL-1α	ND	0 4±0 1	ND
Human dermal fibroblasts	nIL-1β	1 8±0 1	3 4±0 6	2 4±0 1
Human dermal fibroblasts	rIL-1α	1 6±0 1	4 1±0 4	2 8±0 1
BALB-c/3T3 fibroblasts	nIL-1β	2 1±0 6	2 3±1 1	4 8±0 5
BALB-c/3T3 fibroblasts	rIL-1α	3 0±0 5	2 3±1 1	5 5±2 0

LBRM-33-1A5 murine T lymphoma cells[2] and ARH77 human B lymphoma cells were grown in 175 cm^{2} tissue culture flasks in RPMI 1640 containing 10% fetal calf serum (Hyclone) Human dermal fibroblasts were obtained from the American Type Tissue Culture Collection (CRL 1229, passage 5) and propagated in monolayer cultures in DMEM 10% fetal calf serum, cells were tested for IL-1 binding at passage 10-12 BALB-c/3T3 cells were obtained from the American Type Tissue Culture Collection and propagated in DMEM with 10% calf serum (Hyclone) Fibroblasts were harvested for binding assays by treatment with PBS containing EDTA (0 5 mM) to avoid proteolysis of receptor Equilibrium binding data were analysed using a simple bimolecular binding model[12] The values for nonspecific binding were determined by fitting a straight line passing through the origin to data obtained from binding of ^{125}I-ligand to the cells in the presence of 100-fold molar excess of unlabelled IL-1 of the same type These values were subtracted from the data before analysis for determination of K_A and receptors per cell values Inhibition data were analysed with an equation describing competition between two ligands for one class of sites[19] All data analysis was done using RS/1, a commercially available scientific programming package running on a VAX 11/750 under the VMS operating system (BBN Software Products) Error values are standard errors of the mean generated from the nonlinear least-squares fitting program

saturation binding data from these cells for ^{125}I-labelled rIL-1α because the affinity of this ligand for the cells is low (see Fig 2)

To test directly whether IL-1α and IL-1β bind to the same receptor we conducted inhibition experiments (Fig 2) in which the binding of each radiolabelled IL-1 preparation to cells was inhibited by unlabelled IL-1 from the same stock solution and from the other two unlabelled IL-1 preparations For each ^{125}I-labelled IL-1 preparation on either LBRM-33-1A5 or ARH77 cells, complete inhibition of radiolabelled ligand binding was achieved with all three unlabelled IL-1 preparations Thus no receptors on either cell type bound IL-1α or IL-1β exclusively The data in Fig 2 were analysed quantitatively by a nonlinear least-squares fitting of an equation describing competition between labelled and unlabelled hormone for a single class of sites[12,19] The parameter values from this analysis are given in Table 1

The dose-response curves in Fig 2 provide further support for the view that both cell types have a single class of IL-1 receptors that bind all three IL-1 preparations If there existed a subset of receptors that bound one type of IL-1 preferentially, competition between the labelled and unlabelled forms of this type of IL-1 would show a shift to lower concentration of inhibitor than when the unlabelled form competed against labelled heterologous IL-1 This is not observed Finally, for all 18 data sets in Fig 3 the curves pass close to most of the data points A model incorporating only one type of cell surface receptor site thus provides a satisfactory description of the data in all cases

The results of similar experiments with human and murine fibroblasts are summarized in Table 1 Both for 3T3 cells and for human dermal fibroblasts, cross competition experiments revealed that IL-1α and IL-1β bind to identical receptor populations (data not shown) Interestingly, the pattern of relative reactivity differed from one cell line to another, IL-1α bound more avidly than IL-1β to murine T cells but less avidly to human B cells, both forms bound equally well to fibroblasts, irrespective of the species of origin The data suggest that the receptors on different cell lineages, although similar, may not be identical Overall, however, our data are consistent with the view that there are only limited differences in receptor binding activities of IL-1α and IL-1β, and almost no difference in receptor binding between nIL-1β and rIL-1β

The observation that the two forms of IL-1 bind to the same receptor on a variety of cells explains their shared biological activities It is striking that this should occur despite the limited sequence homology between the molecules This may indicate that the tertiary structures are less divergent than the primary sequences, particularly in the active site region This feature of the IL-1 system, that IL-1α and IL-1β bind to the same receptor protein, is also found for the TNF system[20] and the A chain and B chain dimers of PDGF[21] These three hormone-receptor systems appear to have at least partially overlapping roles in controlling connective tissue metabolism and inflammatory responses It is therefore tempting to speculate that the existence of two hormones with a common receptor in all three cases is a consequence of some as yet undiscovered feature of the regulation of responses to disease and injury

Received 16 May, accepted 10 September 1986

1 Gery, I & Waksman, B H *J exp Med* **136**, 143–155 (1972)
2 Gillis, S & Mizel, S B *Proc natn Acad Sci USA* **78**, 1133–1137 (1981)
3 Dewhirst, F E , Stashenko, P P , Mole, J E & Tsurumachi, T *J Immun* **135**, 2562–2568 (1985)
4 Schmidt, J A , Mizel, S B , Cohen, D & Green, I J *Immunology* **118**, 2177–2182 (1982)
5 Dinarello, C A *Rev infect Dis* **6**, 51–95 (1984)
6 Sakatvala, J , Sarsfield, S J & Townsend, Y *J exp Med* **162**, 1208–1227 (1985)
7 Dinarello, C A , Goldin, N P & Wolff, S M *J exp Med* **139**, 1369–1381 (1974)
8 March, C J *et al Nature* **315**, 641–646 (1985)

9 Auron, P E *et al* *Proc. natn. Acad. Sci U S A* **81**, 7907-7911 (1984)
10 Lomedico, P T *et al.* *Nature* **312**, 458-462 (1984)
11 Furutani, Y *et al.* *Nucleic Acids Res* **13**, 5869-5882 (1985)
12 Dower, S K *et al.* *J exp Med.* **162**, 501-515 (1985)
13 Dower, S K, Call, S, Gills, S & Urdal, D L *Proc. natn Acad Sci U S A.* **83**, 1060-1065 (1986)
14 Giri, J G, Lomedico, P T & Mizel, S B *J Immun.* **134**, 343-349 (1985)
15 Kronheim, S R. *et al.* *J exp Med.* **161**, 490-502 (1985)
16 Schmidt, J A *J exp Med* **160**, 772-787 (1984)
17 Segal, D M & Hurwitz, E *J Immun.* **118**, 1338-1347 (1977)
18 Conlon, P J *J Immun* **131**, 1280-1282 (1983)
19 Cuatrecasas, P & Hollenberg, M D *Adv Protein Chem.* **30**, 251-451 (1975)
20 Aggarwal B B, Eessalau, T E & Haas, P E *Nature* **318**, 665-667 (1985)
21 Betsholz, C *et al.* *Nature* **320**, 695-699 (1986)

Presence of Ti (WT31) negative T lymphocytes in normal blood and thymus

Lewis L. Lanier* & Arthur Weiss†

* Becton Dickinson Monoclonal Center, Inc, Mountain View, California 94043, USA
† Department of Medicine, Howard Hughes Medical Institute, University of California, San Francisco, California 94143, USA

The antigen receptor expressed on most T lymphocytes is a disulphide-linked heterodimer (Ti) that is composed of α-chain and β-chain subunits[1-6]. On the surface of human T lymphocytes, Ti is non-covalently associated with three invariant proteins, designated CD3-γ, -δ, and -ε[7]. It has been suggested that Ti is obligatory for CD3 expression[8,9]. But a T leukaemia cell line[10], IL-2 (interleukin 2) dependent T-cell clones established from fetal blood[11] and IL-2 dependent cell lines established from immunodeficiency patients with bare lymphocyte syndrome and ectodermal dysplasia syndrome[12] have recently been shown to express CD3, but not Ti[13] (detected due to monoclonal antibody WT31). These lymphocytes may express the product of the T-cell antigen receptor γ (TCR-γ) gene, rather than the α/β heterodimer, in association with CD3[10,12]. Preliminary studies suggested that T cells expressing CD3 but lacking Ti are present in low frequency in normal lymphoid tissues[10,12]. Here we show that in normal blood and thymus CD3$^+$, WT31$^-$ T cells express neither CD4 nor CD8. The low frequency (<0.2–0 9% of total thymocytes) of CD3$^+$, WT31$^-$ cells in the thymus suggests that this population does not represent a major stage of thymic development and may be a distinct lineage of T cells

Peripheral blood lymphocytes (PBL) were stained by direct two-colour immunofluorescence with fluorescein isothiocyanate (FITC) conjugated anti-Ti (WT31) and biotin conjugated anti-CD3 (Leu 4), followed by phycoerythrin (PE) conjugated streptavidin and analysed by flow cytometry Although most CD3$^+$ T cells co-expressed Ti (WT31), a small, distinct subset of CD3$^+$ T cells did not Based on analysis of >30 normal adult blood donors, ~3% of T lymphocytes were CD3$^+$, WT31$^-$ (range <0 5%-10%), whereas the remainder were CD3$^+$, WT31$^+$ CD3 is expressed on all peripheral blood T cells, but these cells can be divided into subpopulations on the basis of expression of other surface antigens For example, the CD4 and CD8 antigens are mutually exclusive markers for the two major subpopulations of mature T cells This phenotypic heterogeneity has some functional correlation, CD4-bearing cells recognize antigen in association with class II major histocompatibility (MHC) antigens and generally function as helper cells, whereas the CD8 bearing cells recognize antigen in association with class I MHC antigens and generally function as cytotoxic T cells Thus it was of interest to determine whether CD3$^+$, WT31$^-$ cells could be detected in one of these major subpopulations, as this would have implications for cells bearing this unusual CD3 receptor complex. Three-colour immunofluorescence was performed to determine whether CD3$^+$, WT31$^-$ T lymphocytes in blood co-express the CD4 or CD8 antigens Surprisingly, they express neither CD4 nor CD8 (Fig 1) By setting an electronic gate on CD3$^+$ lymphocytes (stained with allophycocyanin (APC)-conjugated streptavidin/biotin-conjugated anti-CD3), it was possible to examine directly CD3$^+$ T cells co-stained with FITC-conjugated anti-Ti (WT31) and a mixture of PE-conjugated anti-CD4 plus PE-conjugated anti-CD8 (Fig 1, right panel) Note that (1) all CD3$^+$, WT31$^-$ T lymphocytes lacked CD4 and CD8, and (2) all detectable CD3$^+$, WT31$^+$ T cells expressed either CD4 or CD8 Identical results were obtained from several independent donors

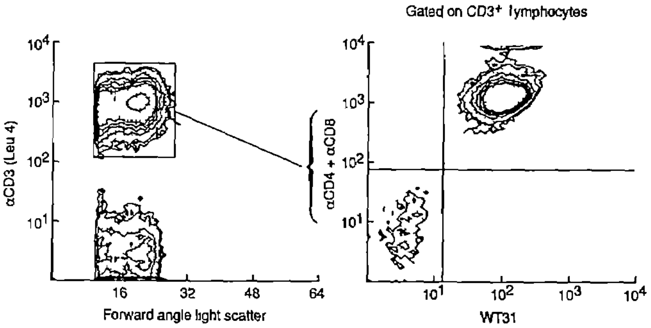


Fig 1 Presence of CD3$^+$, WT31$^-$ T lymphocytes in peripheral blood Mononuclear cells were isolated from peripheral blood using Ficoll/Hypaque Adherent cells were removed using plastic tissue culture flasks PBL were stained with FITC-conjugated WT31, biotin-conjugated anti-CD3 (Leu 4) and a mixture of PE-conjugated anti-CD4 (Leu 3) and anti-CD8 (Leu 2) APC-conjugated streptavidin was used as the second reagent Fluorochrome-conjugated, isotype-matched monoclonal antibodies that do not react with human cells were used to control for nonspecific immunofluorescence Samples were analysed by multi-parameter flow cytometry using a dual laser FACS 440 system with an argon ion laser (488 nm, 100 mW) and a helium neon laser (633 nm, 33 mW) (Becton Dickinson Immunocytometry Systems) Immunofluorescence, flow cytometry and data analysis were performed as described[14,15] Based on the immunoglobulin control sample, contour plots were divided into quadrants to identify unstained cells (lower left), cells stained with both fluorochromes (upper right), and cells stained only with a single fluorochrome (upper left and lower right) In the control sample >99% of lymphocytes were unstained (lower left quadrant, not shown) Left, forward angle light scatter (linear scale, proportional to cell size for lymphoid cells) versus APC anti-CD3 fluorescence Insert, An 'electronic' gate on CD3$^+$ T lymphocytes The correlation of the WT31, CD4 and CD8 antigens on CD3$^+$T lymphocytes is displayed (right panel) In this individual 8% of total lymphocytes were CD3$^+$WT31$^-$ and all lacked CD4 and/or CD8 expression

Further experiments were conducted to determine whether this unique subset of CD3$^+$, WT31$^-$ T cells expressed other T-cell-associated differentiation antigens As shown in Fig 2, >95% of CD3$^+$, WT31$^-$ T lymphocytes co-expressed CD2, CD5 and CD7, similar to expression of these antigens on CD3$^+$, WT31$^-$ T lymphocytes CD1 (a class I MHC homologous antigen limited to thymocytes), CD19 (a pan B cell associated antigen), CD14 (a monocyte associated antigen), transferrin receptors and IL-2 receptors were not expressed on CD3$^+$, WT31$^-$ T cells (not shown)

We have shown previously[10] that CD3$^+$, WT31$^-$ cells comprise <1% of total thymocytes As blood CD3$^+$, WT31$^-$ T cells are CD4$^-$8$^-$, we examined CD4$^-$8$^-$ thymocytes (6% of total thymocytes) for expression of CD3 or WT31 As shown in Fig 3 (top), essentially all thymocytes expressing CD4 and/or CD8 showed linked expression of CD3 and Ti, cells were either CD3$^+$, Ti$^+$ or CD3$^-$, Ti$^-$ The 'diagonal' in the bivariate contour plot indicates a direct quantitative relationship between CD3 and Ti Further analysis by three-colour immunofluorescence

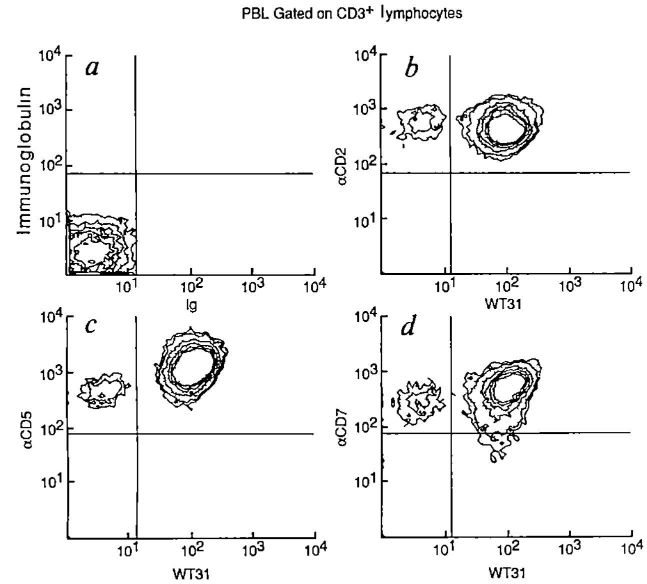


Fig. 2 Expression of T-cell differentiation antigens on CD3$^+$, WT31$^-$ T lymphocytes. PBL (from the same individual as in Fig. 1) were stained with *a*, biotin-conjugated anti-CD3 (Leu 4), FITC-conjugated immunoglobulin control, and PE-conjugated immunoglobulin control, *b*, biotin anti-CD3, FITC-conjugated WT31, and PE-conjugated anti-CD2 (Leu 5) *c*, biotin-conjugated anti-CD3, FITC-conjugated WT31, and PE-conjugated anti-CD5 (Leu 1), *d*, biotin-conjugated anti-CD3, FITC-conjugated WT31, and PE-conjugated anti-CD7 (Leu 9). APC-conjugated streptavidin was used as the second step reagent. An 'electronic' gate was placed on CD3$^+$ T lymphocytes (shown in Fig. 1, left) and the correlation of WT31 with CD2 (B), CD5 (C) or CD7 (D) was directly determined. In control samples (using fluorochrome conjugated isotype-matched immunoglobulin) >99% of lymphocytes were unstained.

(staining with FITC-conjugated anti-CD8, PE-conjugated anti-CD4 and APC-conjugated anti-CD3)[14] indicated that essentially all 'single positive' CD4$^-$8$^+$ and CD4$^+$8$^-$ and about half the immature CD4$^+$8$^+$ thymocytes expressed CD3 (not shown). By contrast, a very minor subset of CD4$^-$8$^-$ thymocytes expressed CD3 but lacked Ti (WT31) (0.8% of total thymocytes, Fig. 3). In fact, all detectable CD3$^+$ lymphocytes in this population were Ti (WT31) negative. Furthermore, like peripheral blood CD3$^+$, WT31$^-$ lymphocytes, CD3$^+$4$^-$8$^-$ thymocytes expressed CD2 and CD5, but not transferrin receptors or IL-2 receptors (Fig. 4). In contrast to peripheral CD3$^+$, WT31$^-$ lymphocytes, CD1 was expressed on a fraction (~40%) of these CD3$^+$, WT31$^-$ thymocytes (Fig. 4*d*), although the amount of CD1 antigen was approximately 10% the amount present on most thymocytes (not shown). Thymocytes from five donors (age 2–6 years) were examined with similar results (mean ~0.5% CD3$^+$, WT31$^-$ thymocytes, range <0.2–0.9%).

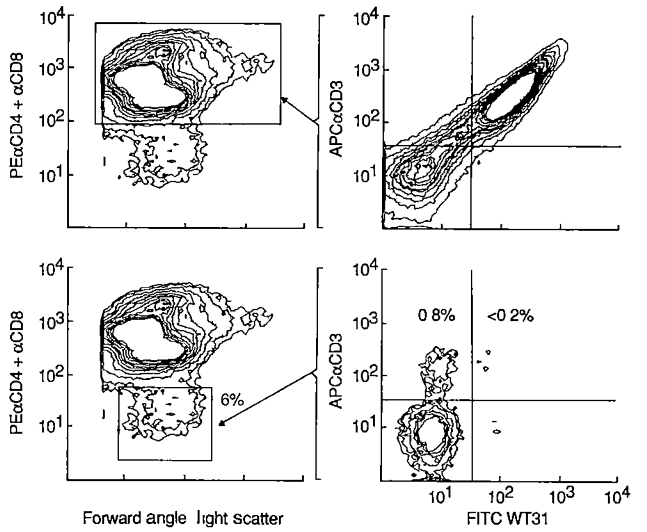
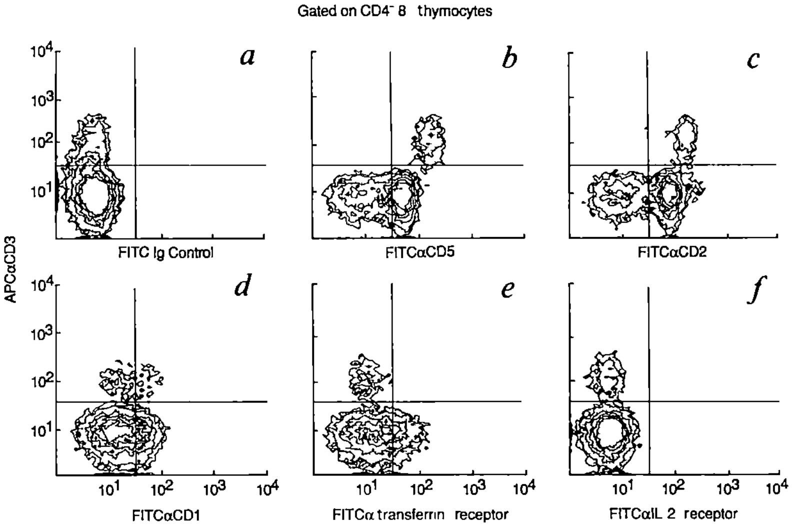


Fig. 3 Expression of CD3 and WT31 on thymocytes. Normal thymus were obtained from paediatric cardiac patients. A single cell suspension of thymocytes was prepared, using the entire thymic lobule to ensure that cortical and medullary thymocytes were accurately represented, and stained with FITC-conjugated WT31, biotin-conjugated anti-CD3 (Leu 4) and a mixture of PE-conjugated anti-CD4 (Leu 3) with anti-CD8 (Leu 2), followed by APC-conjugated streptavidin. An electronic gate was set on thymocytes expressing CD4 and/or CD8 (top) or lacking both CD4 and CD8 (bottom) and the correlation between WT31 and CD3 was determined. In this individual 6% of thymocytes were CD4$^-$8$^-$ and 0.8% were CD3$^+$4$^-$8$^-$, WT31$^-$. In samples stained with fluorochrome-conjugated, isotype-matched control Ig <0.2% of thymocytes were nonspecifically stained (not shown).

We have thus shown that CD3$^+$ T lymphocytes lacking Ti are present in normal blood and at a lower frequency in the thymus. CD3$^+$, WT31$^-$ T cells and thymocytes are only found in the small CD4$^-$8$^-$ subset. Peripheral blood CD3$^+$4$^-$8$^-$ T lymphocytes are morphologically and phenotypically mature, small resting lymphocytes[15]. IL-2 dependent cell lines of CD3$^+$4$^-$8$^-$ thymocytes[16] and normal peripheral blood T lymphocytes[16] established in culture stably express this phenotype, remain Ti

Fig. 4 Expression of T-cell differentiation antigens on CD3$^+$4$^-$8$^-$ thymocytes. Thymocytes (from the same individual as in Fig. 3) were stained with biotin-conjugated anti-CD3 (Leu 4), a mixture of PE-conjugated anti-CD4 (Leu 3) and anti-CD8 (Leu 2), and, *a*, FITC-conjugated Ig control, *b*, anti-CD5 (Leu 1), *c*, anti-CD2 (Leu 5), *d*, anti-CD1 (Leu 6), *e*, anti-transferrin receptor, *f*, anti-IL-2 receptor. APC-conjugated streptavidin was used as the second step reagent. As shown in Fig. 3 (bottom panel), an electronic gate was set on CD4$^-$8$^-$ thymocytes and the correlation between expression of CD3 and of the other differentiation antigens was determined.



(WT31) negative, and demonstrate cytotoxic function (refs 16, 17 and LLL, unpublished) Some cell lines with this phenotype[12,16] do not express mature Ti-α or Ti-β messenger RNA but express a novel heterodimer (probably composed of TCR-γ and another unidentified glycoprotein) in association with CD3 The CD3$^+$ WT31$^-$ cells we have observed *in vivo* are CD4$^-$8$^-$, but some IL-2 dependent lines with this phenotype express CD8[11,12] We do not know whether CD3$^+$8$^+$ WT31$^-$ cells are present *in vivo* in numbers we cannot detect (<0 2%) or whether CD8 arises on these cells as a consequence of culture

These observations raise questions regarding the lineage and differentiation pathway of these unique CD4$^-$8$^-$ T lymphocytes As rearrangement and transcription of TCR-γ genes apparently precedes Ti-β and Ti-α[17–21], it has been proposed that the antigen receptor on immature T cells may be composed of TCR-γ gene products before the Ti-α/β heterodimer appears However because CD3$^+$, WT31$^-$ thymocytes are a minor subpopulation of the putative immature CD4$^-$8$^-$ thymocytes, it is unlikely that CD3$^+$, WT31$^-$ T cells represent a major stage in thymic development If selection of antigen/MHC recognition is from a CD3$^+$, WT31$^-$ thymic subset, there should be more of these cells Because most thymocytes (including the abundant immature CD4$^+$8$^+$ subset) express Ti (WT31), selection may occur mainly via the CD3/Ti receptor complex Alternatively, CD3$^+$Ti$^-$ lymphocytes may be a distinct T-cell lineage and never express Ti The fact that CD3$^+$, WT31$^-$ cells in blood are identical in phenotype to the thymocyte subpopulation (except for CD1, found only on thymocytes) supports this possibility Finally, CD3$^+$, WT31$^-$ peripheral blood cells could be precursors of the CD3$^+$, WT31$^-$ thymocytes which might then be important in thymocyte development If so, a cell population equivalent to the CD3$^+$, Ti$^-$ T cells should be present in athymic mice, and might account for the T-cell-mediated functions observed in these animals[22–25] MacDonald *et al*[26] have recently described a subset of Thy1$^+$, L3T4$^-$, Ly2$^-$ cells in athymic mice and Mak[27] has recently reported that nylon wool non-adherent lymphocytes from athymic mice transcribe TCR-γ, but not Ti-α or competent Ti-β mRNA Further studies are necessary to test this hypothesis and establish the relationship between CD3$^+$, Ti$^+$ and CD3$^+$, Ti$^-$ T lymphocytes

We thank Drs Willem Tax and Patrick Kung for providing the WT31 antibody and Dr Jim Allison for discussion and review of the manuscript

Received 1 August, accepted 8 September 1986

1 Allison, J P, McIntyre, B W & Bloch, D *J Immun* **129,** 2293–2300 (1982)
2 Haskins, K *et al J exp Med.* **157,** 1149–1169 (1983)
3 Meuer, S C *et al J exp Med* **157,** 705–719 (1983)
4 Acuto, O *et al Cell* **34,** 717–725 (1983)
5 Kapler, J *et al. Cell* **35,** 295–302 (1983)
6 McIntyre, B & Allison, J P *Cell* **38,** 659–665 (1984)
7 Borst, J, Prendiville, M A. & Terhorst, C *J Immun* **128,** 1560–1565 (1982)
8 Weiss, A & Stobo, J D *J exp Med* **160,** 1284–1299 (1984)
9 Ohashi, P S *et al Nature* **316,** 606–609 (1985)
10 Weiss, A, Newton, M & Crommie, D *Proc natn. Acad. Sci. U S A* **83,** 6998–7002 (1986)
11 Nowill, A *et al. J exp Med* **163,** 1601–1606 (1986)
12 Brenner, M B *et al Nature* **322,** 145–149 (1986)
13 Oettgen H C, Kappler, J, Tax, W J H & Terhorst, C *J biol. Chem.* **259,** 12039–12048 (1984)
14 Lanier L L Allison, J P & Phillips, J H *J Immun* **137,** 2501–2507 (1986)
15 Lanier, L L, Ruitenberg, J J & Phillips, J H *J exp Med* **164,** 339–344 (1986)
16 Bank, I *et al Nature* **322,** 179–181 (1986)
17 Snodgrass, H R, Dembic, Z, Steinmetz, M & von Boehmer, H *Nature* **315,** 232–233 (1985)
18 Raulet, D H, Garman, R D, Saito, H & Tonegawa, S *Nature* **314,** 103–107 (1985)
19 Garman, R D, Doherty, P J & Raulet, D H *Cell* **45,** 733–742 (1986)
20 Haars R *et al. J exp Med* **164,** 1–24 (1986)
21 Sangster, R N, Minowada, J, Sucin Foca, N, Minden, M & Mak, T W *J exp Med* **163,** 1491–1508 (1986)
22 Hunig, T *Immun Today* **4,** 84–87 (1983)
23 Gillis, S, Union, N A, Baker, P E & Smith, K A *J exp Med* **149,** 1460–1476 (1979)
24 Hunig, T & Bevan, M J *J exp Med* **152,** 688–702 (1980)
25 MacDonald, H R & Lees, R K *J Immun* **132,** 605–610 (1984)
26 MacDonald, H R, Blanc, C, Lees, R K & Sordat, B *J Immun* **136,** 4337–4339 (1986)
27 Mak, T., Yoshikei, Y & Reis, M D *Nature* (in the press)
28 Parks, D R, Lanier, L L & Herzenberg, L A in *Handbook of Experimental Immunology,* 4th edn (eds Weir, D M *et al.*) 29 1–29 2 (Blackwell, Oxford, 1986)

Structure of pre-pro-von Willebrand factor and its expression in heterologous cells

David T. Bonthron*†, Robert I. Handin‡, Randal J. Kaufman§, Louise C. Wasley§, Elizabeth C. Orr§, Lisa M. Mitsock§, Bruce Ewenstein‡, Joseph Loscalzo‡, David Ginsburg*‡|| & Stuart H. Orkin*†

* Division of Hematology/Oncology, Children's Hospital,
300 Longwood Avenue, Boston, Massachusetts 02115,
USA and the Dana-Farber Cancer Institute, 44 Binney Street,
Boston, Massachusetts 02115, USA
† Howard Hughes Medical Institute and the Departments of
Medicine and Pediatrics, Harvard Medical School,
25 Shattuck Street, Boston, Massachusetts 02115, USA
‡ Division of Hematology, Brigham and Women's Hospital,
75 Francis Street, Boston, Massachusetts 02115, USA
§ Genetics Institute, 87 Cambridge Park Drive, Cambridge,
Massachusetts, USA

Von Willebrand factor (vWF), a multifunctional haemostatic glycoprotein derived from endothelial cells and megakaryocytes, mediates platelet adhesion to injured subendothelium and binds coagulation factor VIII in the circulation. Native vWF is a disulphide-bonded homopolymer; the monomeric subunits, of apparent relative molecular mass (M_r) 220,000 (220K) are derived from an intracellular precursor estimated at 260–275K (refs 1–5). Multimer assembly is preceded by the formation of dimers, linked near their C-termini, which then assemble into filamentous polymers. The importance of the removal of the large vWF pro-polypeptide during multimer assembly, and whether this or other stages of the complex post-translational processing require components specific to endothelial cells or megakaryocytes, is unknown. Here we report an analysis of the complete sequence of pre-pro-vWF and expression of the molecule in heterologous cells. The vWF precursor is composed of several repeated subdomains. When expressed in COS and CHO cells, it is cleaved and assembled into biologically active high relative molecular mass disulphide bonded multimers. This suggests that the information for assembly of this complex molecule resides largely within its primary structure.

The nucleotide sequence of pre-pro-vWF complementary DNA (which is printed in its entirety elsewhere[48]) has a major open reading frame encoding 2,813 amino acids The initiator methionine, followed by a classical signal peptide, can be identified unambiguously because it is the only Met between an in-frame upstream stop codon and the Ala (residue 23) determined by amino-acid sequencing to be the N-terminus of the 100K pro-polypeptide[6] The signal peptidase cleavage site (defined by this N-terminal sequence) conforms well to the consensus[7] The predicted M_r of pro-vWF is 307K without the extra 15% contributed by glycosylation[8], this is considerably larger than the size of the intracellular precursor estimated from SDS-PAGE[5]

The amino-acid sequence of pre-pro-vWF shows extensive internal repetition (Fig 1) A central region of ~600 amino-acids contains three domain A repeats, which together with the short B and C repeats, were identified in partial-length clones described by others[9,10] A striking feature of the A region is its lack of cysteine residues, in contrast, the flanking regions (amino acids 34–1,242 and 1,948 to the C-terminus) are rich in cysteine, the most abundant amino acid in pre-pro-vWF, comprising 8 3% of the residues (10 4% in the regions outside domain A) These cysteine-rich regions include a further series of four repeats, which we term D1–D4, beginning just after the signal peptide (Fig 1) The second of these 350 amino-acid repeats is truncated,

---

|| Present address Howard Hughes Medical Institute, 1150 West Campus Drive, Ann Arbor, Michigan 48109, USA

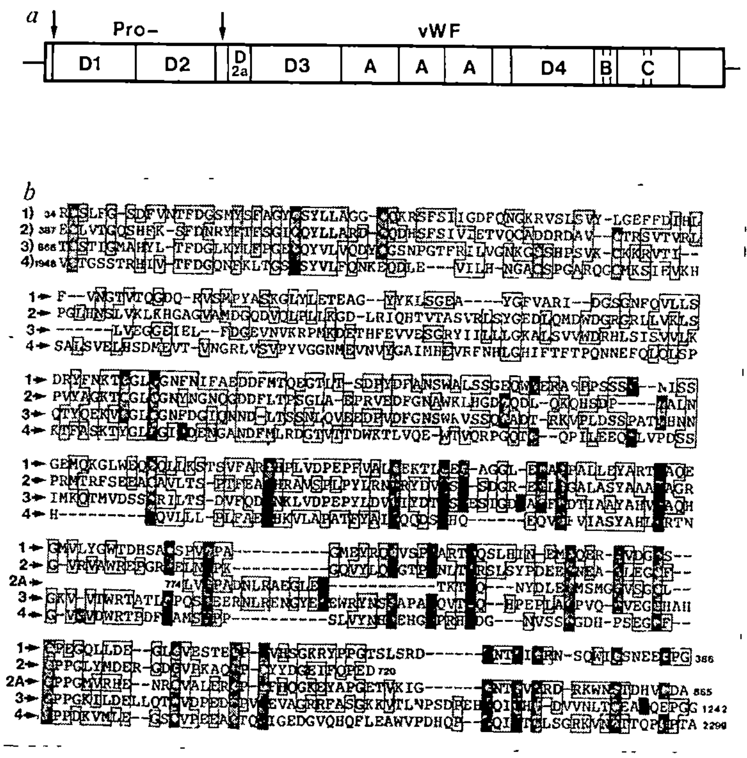


Fig. 1 Repeated elements in the primary structure of vWF *a*, Schematic of vWF cDNA D1–D4, cysteine-rich repeat blocks The A, B and C repeat domains of Sadler *et al*[9,10] are also labelled Arrows, sites at which the signal peptide and the large pro-polypeptide are cleaved from the primary translation product *b*, Alignment of the four cysteine-rich D repeat blocks Matching residues are boxed and lightly shaded, cysteine residues are darkly shaded Small numbers at the sides refer to the adjacent amino-acid residues, relative to the initiator methionine, assigned position 1 Amino acids 721–773, (between repeats 2 and 2A), and 1,243–1,947 (including the A domain repeats), have been omitted, the sequence is otherwise continuous as shown Gaps introduced to maximize homology are shown as dashes D4 appears to be the least conserved of the repeat blocks, in this alignment, the percentage of matching amino acids for each pairwise comparison is D1–D2, 34 8, D1–D3, 30 4, D2–D3, 33 6, D1–D4, 24 1, D2–D4, 25 0, D3–D4, 24 1 The sequence TCGLC, which forms the centre of a highly conserved block in the D repeat regions, also occurs in the B domain, indicating a possible relationship between these two repeated elements

Methods The amino-acid sequence shown is deduced from the sequences of the same cDNA clones used in construction of pvWF2 (Fig 2 legend) Both strands of the sequence were determined by the Sanger dideoxy method[34,35] M13 subclones for sequencing were constructed by exonuclease digestion methods[36,37] The region of residue 1,472 (see text) was sequenced in three further independent clones (pvWG8b, pvWH22, pvWH23) using an internal oligonucleotide primer[38] Repeat blocks were defined manually following identification of short repeated motifs by computer search The consensus alignment was also arrived at by hand, with reference to computer-generated pairwise alignments

and its C-terminal one-third duplicated (Fig 1, D2a) Interestingly, the site of pro-polypeptide cleavage (Fig 1, arrow) falls within this region of discontinuity of the D repeats

Based on partial cDNA sequences, Sadler and colleagues[10] proposed the existence of a fifth repeated element which they term domain E Our analysis shows that this element is, in fact, the C-terminal portion of the D repeat block, and these sequences are included in the alignment in Fig 1

Within the D repeats there is striking conservation of the positions of the abundant cysteine residues (Fig 1*b*) Because no free sulphydryl groups can be detected in native vWF[11,12] all these cysteines must participate in inter- and intra-chain disulphide bonding, an important determinant of the tertiary and quaternary structure of the protein Electron microscopy of vWF shows a central filament separating globular regions at each end of the monomer[13–15] These two structural features may correspond to the cysteine-poor A domains and the highly disulphide-bonded cysteine-rich domains, respectively The intron-exon arrangement underlying the repeated structure of vWF has not yet been defined, but the common evolutionary origin of mature vWF and the pro-polypeptide is evident from our analysis As the cleavage of pro-vWF occurs at a position where the cysteine-rich repeat sequence is interrupted by a partial duplication, it could be that after the duplications of the D repeats a later aberrant recombination introduced the new cleavage site, leading to the separation of the pro-piece from mature vWF Despite the functional similarities between vWF and fibronectin[16] and fibrinogen[17], we find no similarity in sequence between these proteins other than the consensus tetrapeptide Arg-Gly-Asp-Ser, which defines the binding site for platelet glycoprotein IIb/IIIa and is common to many adhesive glycoproteins[18–20]

Discrepancies exist between our full-length sequence and the partial sequence of Sadler *et al*[9,10] The first 13 base pairs (bp) of the clone λHvWF-1 (ref 9) differ completely from our sequence, which in this region was determined on two independent clones, pvWH33 and pvWH5 At amino-acid residue 770 (residue 7 of the mature monomer), the Sadler *et al* sequence has His rather than the Pro determined from N-terminal amino-acid sequencing[9], whereas our clones H33 and H5 both have the expected Pro At residue 1,472, our sequence contains CAC (His) rather than GAC (Asp) Of three further non-identical clones from our endothelial cDNA library, two have GAC and one CAC at this position, suggesting an amino-acid polymorphism Some other potential polymorphisms have been previously noted[10]

To study vWF biosynthesis in heterologous cells, we assembled a full-length cDNA and inserted it into the expression vector pMT2 (Fig 2 legend), in which transcription is directed by the adenovirus major late promoter In assays of transient expression, the conditioned medium of monkey kidney (COS) cells transfected with the pMT2-vWF construct contained 50–100 ng ml^{-1} of vWF antigen COS cells transfected with pMT2 alone did not produce material reacting in our inhibition enzyme-linked immunosorbent assay (ELISA)

For expression of vWF in Chinese hamster ovary (CHO) cells, we used a second expression vector, pMT2ADA-vWF (Fig 2) with a protocol of selection for cells over-expressing the enzyme

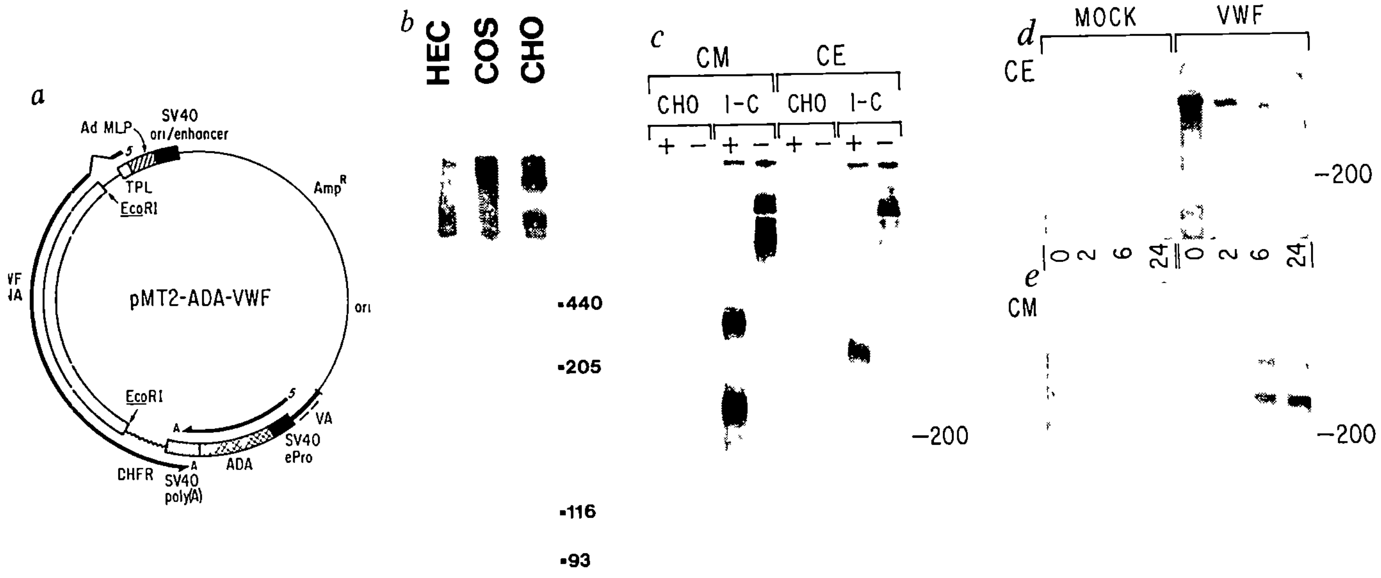


Fig 2 Expression of vWF cDNA *a*, Plasmid pMT2ADA-vWF used for expression in CHO cells It is similar to pMT2-vWF, but contains in addition a human adenosine deaminase (ADA) transcription unit AdMLP, adenovirus major late promoter, TPL, tripartite leader VA genes ePro, simian virus 40 (SV40) early promoter *b–e*, Analysis of immunoprecipitated material from cell extracts or conditioned medium of passaged human endothelial cells (HEC), COS cells or CHO cells *b*, Immunoprecipitates from conditioned medium analysed unreduced on a 4 3% gel (extended run time) *c*, Labelled immunoprecipitates from conditioned medium (CM) or cell extracts (CE) of CHO cells (untransfected) and the transfected, ADA-selected line 1−C, with (+) or without (−) reducing agent *d*, Pulse-chase of ^{35}S-labelled material from transfected or mock-transfected COS cell extracts *e*, Pulse-chased material in conditioned medium from the same experiment

Methods Segments of overlapping cDNA clones[32] were recombined as follows pvWH33 from the 5′ *Eco*RI linker to the unique *Bam*HI site, pvWE6 from the *Sac*II site to the 3 *Eco*RI linker, pvWH5 between the above *Bam*HI and *Sac*II sites The full-length cDNA in plasmid pUC13[39] was designated pvWF2 For transient expression in COS cells, the *Eco*RI fragment containing the entire coding region was transferred to the expression vector pMT2, generating pMT2-vWF pMT2 is a derivative of the mammalian cell expression vector p91023(B)[40] in which the ampicillin resistance marker has replaced the tetracycline resistance marker pMT2-vWF was grown in *Escherichia coli* DH5 (*recA*) to prevent deletion of vWF sequences The expression vector used for CHO cells was constructed by inserting a *Pvu*II/*Hin*fI fragment from pSV2 ADA (Ia)[41], containing the SV40 early promoter and human ADA coding region, into a unique *Xba*I site in pMT2 This generated a transcription unit that uses the SV40 late polyadenylation signal present in pMT2 for expression of ADA The vWF cDNA fragment was then inserted at the unique *Eco*RI cloning site to generate pMT2ADA-vWF (*a*) Protoplasts of *E coli* DH5 carrying pMT2ADA-vWF were fused to CHO cells as described[42] After 2 days, the cells were subcultured into selective medium containing 1 1 mM adenosine, 1 mM uridine, 50 μM alanosine and 0 1 μM 2′ deoxycoformycin[20] Transformants (~100) were pooled and propagated sequentially in 0 5 and 1 0 μM 2′ deoxycoformycin to select for amplification of the ADA and vWF cDNAs One such line obtained was designated 1−C For assay of vWF, these cells were rinsed and incubated for 24 hours in serum-free medium COS cells were transfected with pMT2-vWF or with pMT2 by the DEAE-dextran technique, with chloroquine treatment[43–45] Forty-eight hours after transfection, cells were rinsed and fed serum-free medium for 24 h Culture medium was then collected vWF antigen was measured by an inhibition ELISA, using affinity-purified rabbit anti-vWF antiserum, purified vWF antigen and biotinylated second antibody (goat anti-rabbit IgG) with colour development by avidin–peroxidase complex COS cells were labelled with ^{35}S-cysteine (0 5 mCi ml^{-1}, NEN) for 15 min, 60 h after transfection They were then fed with serum-free medium containing cysteine, and incubated at 37 °C for the times indicated Control CHO cells and the 2′ deoxycoformycin-resistant CHO cells were labelled for 3 h with ^{35}S-methionine (0 5 mCi ml^{-1}), followed by addition of an equal amount of label for a further 3 h CM and CE were prepared for immunoprecipitation as described[46] Immunoprecipitation was carried out with a 1 40 dilution of affinity purified rabbit anti-vWF antibody for 8–12 h at room temperature, followed by either Staphylococcal protein A or goat anti-rabbit IgG immobilized on Sepharose beads After 2 h incubation the beads were washed and taken up in SDS-gel loading buffer with or without 1 mM dithiothreitol, heated to 100 °C and run on 6% (*d–e*) or 4 3% (*b–c*) polyacrylamide gels using the Laemmli buffers[47], followed by drying and autoradiography

adenosine deaminase to amplify the plasmid sequences[21] The cell line obtained (1−C) secretes 4 μg ml^{-1} vWF antigen per 10^6 cells per 24 h as measured by ELISA, a level approaching that of normal plasma (5–10 μg ml^{-1})

Functional vWF multimers are normally assembled only in endothelial cells and megakaryocytes via a complex series of processing steps[3–5] The vWF pro-polypeptide cleavage site, preceded by the basic amino acids Lys-Arg, is similar to the cleavage sites of many prohormones (see, for example, refs 22–27) Although COS cells can correctly process at least one such site[28], the requirements for precursor cleavage have not been defined, and correct removal of the large vWF pro-polypeptide could not be confidently expected Furthermore, the assembly of vWF monomers into disulphide-bonded dimers and subsequently into linear arrays of high M_r multimers might require enzymes or organelles restricted to the specialized cell types that normally secrete vWF

To investigate the post-translational processing of recombinant vWF, we labelled cells with ^{35}S Cys or Met and immunoprecipitated vWF from medium and cell extracts We analysed unreduced immunoprecipitates on 4 3% polyacrylamide gels A series of high M_r bands is seen in medium from passaged human endothelial cells (HEC) and from transfected COS and CHO cells (Fig 2*b*), although the lack of resolution in this region of the gel precludes detailed comparison of the multimer structure On reduction and denaturation, the immunoprecipitated material from the medium (CM) of transfected CHO cells (Fig 2*c*) or COS cells (not shown) runs as the expected two bands of M_r 260K and 220K Control immunoprecipitates from untransfected CHO cells are negative (Fig 2*c*, lanes 1, 2) Cell lysates (CE) contain predominantly the precursor species (Fig 2*c*, lane 7), which appears slightly smaller than the precursor seen in conditioned medium, in accord with previous observations on endothelial cells[5] In pulse-chase experiments on transfected COS cells (Fig 2*d*, *e*, cell extract and conditioned medium, respectively), labelled vWF, predominantly in the processed form, is present in the cell medium from 6 h, while the labelled precursor disappears from the cell extracts over a similar time period

These studies indicate that cells expressing recombinant vWF not only remove the vWF pro-polypeptide but can assemble high M_r disulphide-bonded multimers from the cleaved product To assess the functional properties of recombinant vWF, we examined its interactions with platelet glycoprotein Ib and with collagen Binding of native vWF multimers to these two macromolecules results in attachment of platelets to vascular

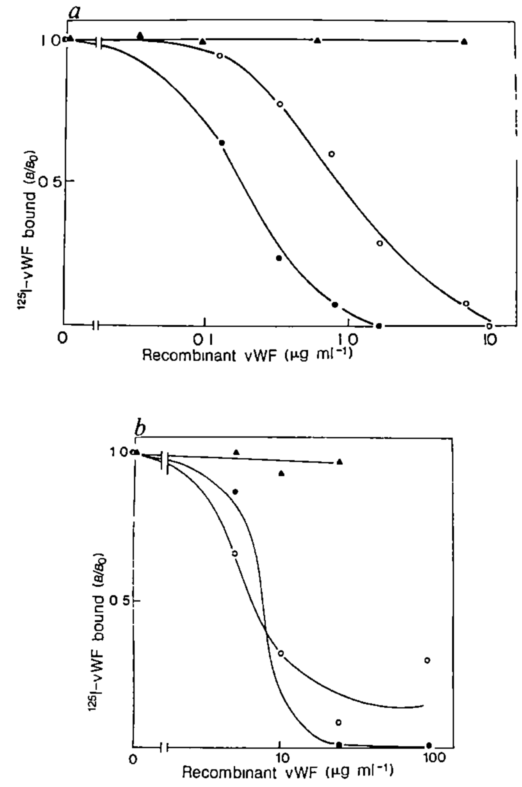


Fig 3 Biological activity of recombinant vWF The ability of culture media derived from COS cells transfected with pMT2 or pMT2-vWF to compete with purified native vWF for binding to platelet glycoprotein Ib (GpIb) or to collagen coated monolayers is shown *a*, Competition for binding to collagen-coated monolayers (closed circles) is compared with competition by purified plasma vWF (open circles) *b*, Competition for binding to GpIb using ristocetin to induce binding to freshly prepared washed platelets (closed circles) or formalin-fixed platelets (open circles) The control supernatants did not compete in either experiment (triangles)

Methods vWF multimers were purified from normal human plasma and iodinated by the Iodobead technique[29] To raise the concentration of recombinant vWF from COS cells sufficiently to carry out the radioligand assays, 50-100 ml conditioned medium was dialysed against 50% Ficoll for 6 to 8 h, and the 10 fold-concentrated samples were reassayed with the inhibition ELISA For analysis of GPIb binding, platelets were washed free of plasma and used immediately or treated with 1% paraformaldehyde and stored at 4 °C for several weeks Binding of ^{125}I-vWF was induced by the addition of 1 5 mg ml^{-1} final concentration of ristocetin, as previously described[29] For the collagen binding assays, a 1 mg ml^{-1} solution of acid-soluble types I and III collagen at *p*H 3 0 was added to the wells of an Immulon microtitre plate Unbound collagen was aspirated after 1 h, and the wells washed with a series of buffered solutions to raise the *p*H to 7 4 Labelled vWF and competing ligand were added to the wells, and after 1 h at room temperature unbound material was removed by aspiration and several washes with 150 mM NaCl, 10 mM Tris, *p*H 7 6 The wells were then detached from the plate and bound radioactivity determined In previous studies binding was found to be specific, saturable and reversible, and to require vWF multimers and native collagen[30]

subendothelium We used concentrated COS cell supernatants as competing ligand against purified, ^{125}I-labelled human plasma vWF multimers under conditions in which native vWF binds specifically and saturably (Fig 3, refs 29, 30) The concentration of recombinant vWF displacing 50% of native vWF from collagen (IC$_{50}$) is 0 2 μg ml^{-1}, fivefold lower than the IC$_{50}$ obtained using purified human vWF For GP Ib binding, the IC$_{50}$ for recombinant vWF (Fig 3*b*) and plasma vWF (not shown) are equivalent (5-7 μg ml^{-1}) Control medium from cells transfected with the vector pMT2 does not compete (Fig 3*a*, *b*, solid triangles) Multimeric recombinant vWF thus exhibits similar activity to native vWF in two relevant bioassays

The availability of vWF cDNA clones[31-33] permits new experimental approaches to the complex biosynthesis of this protein Here, we show that recombinant vWF is processed and secreted in high M_r form by two heterologous cell types The multimer conformation is sufficiently like that of native vWF that the recombinant material competes effectively for binding to platelets and collagen, In fact, recombinant vWF competes more effectively than vWF derived from human plasma for binding to collagen, possibly because some of the highest M_r and most active multimers are lost during purification of plasma vWF, whereas recombinant vWF was not fractionated or chromatographed before analysis This suggests that assembly of vWF multimers is directed largely by the sequence of the primary translation product rather than by endothelial- or megakaryocyte-specific machinery More detailed analysis of the structure of recombinant vWF multimers is in progress

Our analysis reveals similarities between the mature vWF subunit and the large vWF pro-polypeptide These two related glycoproteins are both secreted after cleavage[6], but possible functions of the pro-polypeptide in the circulation, or in linkage of the N-termini of dimers during multimer assembly, remain undefined Site-directed mutagenesis and expression of the cDNA in cultured cells can now be used to identify features of pro-vWF that are important for its processing and biological activity

This work was partly supported by NIH grants to R I H and S H O We thank Patricia Murtha and Monique Davies for technical help, and Dr Robert Wise for comments and discussion S H O is an Investigator of the Howard Hughes Medical Institute

Received 2 May, accepted 8 August 1986

1 Legaz, M E *et al J biol Chem* **248,** 3946-3955 (1973)
2 Shapiro, G A *et al J clin Invest* **52,** 2198-2210 (1973)
3 Wagner, D D & Marder, V J *J biol Chem* **258,** 2065-2067 (1983)
4 Lynch, D C *et al J biol Chem* **258,** 12757-12760 (1983)
5 Wagner, D D & Marder, V J *J Cell Biol* **99,** 2123-2130 (1984)
6 Fay, P J *et al Science* **232,** 995-998 (1986)
7 Perlman, D & Halvorson, H O *J molec Biol* **167,** 391-409 (1983)
8 Federici, A B *et al J clin Invest* **74,** 2049-2055 (1984)
9 Sadler, J E *et al Proc natn Acad Sci U S A* **82,** 6394-6398 (1985)
10 Shelton Inloes, B B, Titani, K & Sadler, J E *Biochemistry* **25,** 3164-3171 (1986)
11 Counts, R B, Paskell, S L & Elgee, S K *J clin Invest* **62,** 702-709 (1978)
12 Hessel, B *et al Thromb Res* **35,** 637-651 (1984)
13 Ohmori, K *et al J Cell Biol* **95,** 632-640 (1982)
14 Slayter, H *et al J biol Chem* **260,** 8559-8563 (1985)
15 Fowler, W E *et al J clin Invest* **76,** 1491-1500 (1985)
16 Mosher, D F *Prog Hemost Thromb* **5,** 111-151 (1980)
17 Doolittle, R F *Scient Am* **245,** 126-130 (1981)
18 Pierschbacher, M D & Ruoslahti, E *Nature* **309,** 30-33 (1984)
19 Pytela, R *et al Science* **231,** 1559-1562 (1986)
20 Ruoslahti, E & Pierschbacher, M D *Cell* **44,** 517-518 (1986)
21 Kaufman, R J *et al Proc natn Acad Sci U S A* **83,** 3136-3140 (1986)
22 Goodman, R H *et al J biol Chem* **257,** 1156-1159 (1982)
23 Vasicek, T J *et al Proc natn Acad Sci U S A* **80,** 2127-2131 (1983)
24 Bell, G I *et al Nature* **304,** 368-371 (1983)
25 Wiborg, O *et al Proc natn Acad Sci U S A* **81,** 1067-1069 (1984)
26 Rosenfeld, M G, Amara, S G & Evans, R M *Science* **225,** 1315-1320 (1984)
27 Leiter, A B *et al J biol Chem* **260,** 13013-13017 (1985)
28 Warren, T G & Shields, D *Cell* **39,** 547-555 (1984)
29 Bockenstedt, P, Greenberg, J M & Handin, R I *J clin Invest* **77,** 743-749 (1986)
30 Bockenstedt, P, McDonagh, J & Handin, R I *J clin Invest* **78,** 551-556 (1986)
31 Lynch, D C *et al Cell* **41,** 49-56 (1985)
32 Ginsberg, D *et al Science* **228,** 1401-1406 (1985)
33 Verweij, C L *et al Nucleic Acids Res* **13,** 4699-4717 (1985)
34 Sanger, F, Nicklen, S & Coulson, A R *Proc natn Acad Sci U S A* **74,** 5463-5467 (1977)
35 Biggin, M D, Gibson, T J & Hong, G F *Proc natn Acad Sci U S A* **80,** 3963-3965 (1983)
36 Poncz, M *et al Proc natn Acad Sci U S A* **79,** 4298-4302 (1982)
37 Dale, R M K, McClure, B A & Houchins, J P *Plasmid* **13,** 31-40 (1985)
38 Chen, E Y & Seeburg, P H *DNA* **4,** 165-170 (1985)
39 Vieira, J & Messing, J *Gene* **19,** 259-268 (1982)
40 Wong, G G *et al Science* **228,** 810-815 (1985)
41 Orkin, S H *et al Molec cell Biol* **5,** 762-767 (1985)
42 Sandri Goldin, R M *et al Molec cell Biol* **1,** 743-752 (1981)
43 Lopata, M A, Cleveland, D W & Sollner-Webb, B *Nucleic Acids Res* **12,** 5707-5717 (1984)
44 Luthman, H & Magnusson, G *Nucleic Acids Res* **11,** 1295-1308 (1983)
45 Kaufman, R J *Proc natn Acad Sci U S A* **82,** 689-693 (1985)
46 Kaufman, R J *et al Molec cell Biol* **5,** 1750-1759 (1985)
47 Laemmli, U K *Nature* **227,** 680-685 (1970)
48 Bonthron, D T *et al Nucleic Acids Res* (in the press)

A nocturnal inhibitor of carboxylation in leaves

S. Gutteridge*§, M. A. J. Parry*, S. Burton*, A. J. Keys*, A. Mudd*, J. Feeney†, J. C. Servaites‡ & J. Pierce§

* Rothamsted Experimental Station, Harpenden, Herts AL5 2JQ, UK
† National Institute for Medical Research, Mill Hill, London NW7 1AA, UK
‡ Department of Biology, University of Dayton, Dayton, Ohio 45469-0001, USA
§ Central Research and Development Department, Experimental Station, E I DuPont Company, Wilmington, Delaware 19898, USA

The diurnal variation in the activity of ribulose-1,5-bisphosphate carboxylase (RuBPCase), the major CO_2-fixing enzyme in plants, has been shown to result from the influx and efflux of Mg^{2+} ions into and out of the chloroplast stroma. A recent re-examination of the phenomenon indicates that the inactivation of the enzyme, rather than being due to the efflux of Mg^{2+}, is correlated in some plant species with an increase in the concentration of an organic phosphate ester in the chloroplast in the dark[1-3]. We have purified this potent inhibitor from dark-treated potato (*Solanum tuberosum*) leaves, and established that its structure is 2-carboxy-D-arabinitol-1-phosphate, a molecule that closely resembles an intermediate in the carboxylase reaction of RuBPCase.

The method used to isolate the inhibitor from dark-treated plant material exploited its affinity for the active enzyme The carboxylase was extracted and purified in buffer containing CO_2 (HCO_3^-) and Mg^{2+} so that the inhibitor remained bound The specific carboxylase activity of the enzyme at the first stages of purification was only 40% that of the enzyme purified from leaves during the day This activity could be more than doubled by treatment with sulphate to release the inhibitor from the enzyme[1] Full details of the purification procedure are given in the legend to Table 1 Figure 1 shows the inhibition of purified wheat RuBPCase as a function of the total phosphate content of the purified inhibitor at various dilutions Maximum inhibition was seen at inhibitor concentrations stoichiometric with those of the enzyme's active site

There are a number of phosphate esters synthesized in the chloroplast[4] (for example, 6-phosphogluconate) that may act as potent inhibitors of the carboxylase The purified inhibitor was compared to such compounds by thin layer chromatography on cellulose powder and in several solvent systems The R_f values of the natural inhibitor and its dephosphorylated product indicated that it was unlike any of the phosphorylated sugars or sugar alcohols usually found in plants

Table 1 Inhibitor purification from dark-harvested potato leaves

Fraction	Weight of dry residue (mg)	Total phosphorus (μg atom)
Extract in 80% methanol	1,500	11 9
Sephadex G-10 column	92	10 6
Polar fraction from C_{18} Sep-Pak	16	7 6
QAE-A25 column	1 7	6 1

To obtain enough material for detailed analysis, 1 kg of potato leaves were collected early in the morning and the soluble proteins extracted at *p*H 8 0 in a buffer containing 20 mM $MgCl_2$ and 20 mM $NaHCO_3$ RuBPCase with the inhibitor bound was isolated from the plant debris as a precipitate using polyethylene glycol The precipitate was redissolved in a minimum volume of the activation buffer, 0 1 M bicine *p*H 8 0 containing 20 mM $MgCl_2$ and 20 mM $NaHCO_3$ Further purification was achieved by chromatography on an S400 (Pharmacia) gel filtration column The inhibitor was released from the enzyme by addition of methanol to give a final concentration of 80% v/v Precipitated protein, 1 3 g, was removed by centrifugation and re-extracted with further additions of 80% methanol The combined extracts were concentrated and desalted into a volatile buffer (10 mM NH_4HCO_3) using a Sephadex G-10 column and those fractions exhibiting the strongest inhibition of wheat RuBPCase were evaporated to dryness This step removed $MgCl_2$ introduced with the buffer used in extraction and purification of the potato enzyme The residue was dissolved in distilled water, adjusted to *p*H 4 0 with formic acid, and further purified by passage through a C_{18} Sep-Pak column (Waters) The final treatment was consecutive ion-exchange chromatography using anionic and cationic columns[9] The resulting compound, dried to constant weight, was suitable for spectrophotometric analysis

A sample containing 2 5 μmol of the inhibitor (based on total phosphate) was dissolved in 0 4 ml of D_2O (*p*D ~ 3 0) and the solution was examined by proton NMR spectroscopy using a Bruker 500 MHz instrument The spectrum consists of a number of multiplets arising from only six protons associated with four C centres Selective irradiation spin decoupling experiments indicate that three of these centres are adjacent to one another Consideration of the spin coupling constants and the chemical shifts for these protons suggests a structure of $HOCH_2$-CH(OH)-CH(OH)- The remaining multiplets correspond to a pair of methylene protons that are not coupled to the other protons in the molecule and have chemical shifts and additional splittings ($J_{^{31}P-^{1}H} = 5$ Hz) characteristic of a $-CH_2OPO_3^{2-}$ group[5] Thus the inhibitor has the general structure CH_2(OH)-CH(OH)-CH(OH)-X-$CH_2OPO_3^{2-}$ The NMR spectrum is *p*H dependent, consistent with the presence of a carboxylic acid group as a branch to the carbon chain, and at low *p*H cyclization to a lactone ring can occur X was therefore considered to be a quaternary carbon with a carboxylic group, that is >C(OH)-CO_2^- This suggested that the compound might be similar to the potent carboxypentitol inhibitors 2-carboxyarabinitol-1,5-bisphosphate and 2-carboxyribitol-1,5-bisphosphate (2CABP and 2CRBP, respectively) but with a single phosphate group at the 1 position

The confirmation that the natural inhibitor must be a 2-carboxy pentitol-1-phosphate was achieved by first dephos-

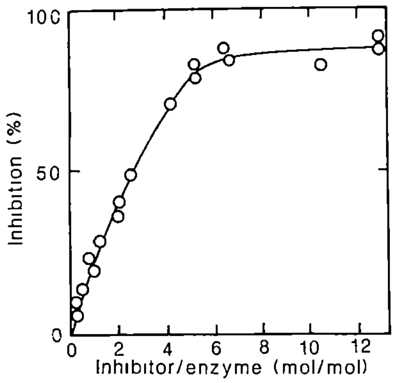


Fig 1 Inhibition of carboxylase activity of purified wheat RuBPCase[8] by the inhibitor from potato leaf The carboxylase was activated (2 mg ml^{-1}) in 0 1 M bicine, *p*H 8 2 at 40 °C for 40 min in the presence of 10 mM $NaHCO_3$ and 20 mM $MgCl_2$ Samples of this solution (20 μl) were mixed with 20 μl of a solution containing the inhibitor After at least 2 min, 20 μl of the mixture was assayed for carboxylase activity at 25 °C The inhibitor concentration is based on the phosphate content of the solutions

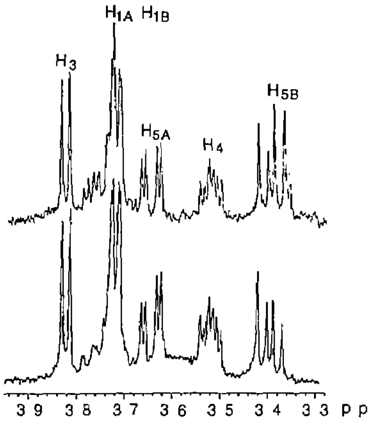


Fig. 2 The proton NMR spectrum at 500 MHz of 2-carboxyarabinitol-1-phosphate. *a*, The natural compound, isolated by dissociation from the active site of RuBPCase and then purified as outlined in the legend to Table 1. *b*, The spectrum of the monophosphate synthesized from 2-carboxyarabinitol-bisphosphate by partial hydrolysis with alkaline phosphatase (S.G. and M.A.J.P., unpublished data). The chemical shifts are given in Table 2. The *p*D of the samples was 8.0 and hence the spectra are for the open chain form.

phorylating the compound with alkaline phosphatase and then producing the trimethyl silylated derivative for examination by gas chromatography mass spectrometry. One of the major silylated products has an ion ($m/z = 613$, M-15$^+$) that is consistent with a fully trimethyl silylated derivative of carboxypentitol. A second ion ($m/z = 466$, M$^+$) was also observed corresponding to the trimethyl silylated lactone form of a carboxypentitol. The same molecular ions were produced when 2-carboxyarabinitol was synthesized from 2CABP by phosphatase treatment and silylated under the same conditions.

There are four epimers that satisfy the general structure described above. The natural compound could be 2-carboxyarabino- or 2-carboxyribo- stereoisomers (epimeric about C_2) or 2-carboxylyxitol or 2-carboxyxylitol monophosphates (epimeric about C_3). To distinguish between these possibilities, the bisphosphates were synthesized as previously described[6,7], the phosphate groups were removed and the proton spectra of the resulting acids compared to the dephosphorylated natural inhibitor. It is quite clear from the chemical shifts of the protons (Table 2) that the compound is identical to 2-carboxy-D-arabinitol, therefore, the natural phosphate ester is 2-carboxy-D-arabinitol-1-phosphate. Its proton spectrum is shown in Fig. 2 along with the spectrum of the monophosphate synthesized from 2CABP.

2-carboxyarabinitol-1-phosphate has many features enabling it to act as a potent and yet reversible inhibitor of RuBPCase. It closely resembles the 2-carboxy-3-ketoarabinitol-1,5-bisphosphate intermediate of the carboxylase reaction (see Fig. 3). However, unlike 2CABP the absence of a second phosphate group on the primary alcohol group at the five position ensures that it does not become bound irreversibly to the active site. The response of RuBPCase to effectors, particularly transition state analogues, in terms of activation and activity *in vitro* have been exhaustively studied (see ref. 4 for a review), but this is the first report of a compound with this type of structure functioning in this way in the chloroplast. The function of the transition state analogue as an inhibitor of RuBPCase *in vivo* is unknown. During normal photosynthesis, the enzyme is activated by carbamylation with CO_2 thus forming the binding site for the essential Mg^{2+} ion. The substrate, ribulose bisphosphate, binds to the active site and the close interaction with the metal initiates the catalytic process[5]. However, if the substrate binds to the active site before the reaction with CO_2 and Mg^{2+} it very effectively blocks the activation process. An inhibitor that can dissociate from the active site and yet is preferentially bound to the activated form of the enzyme, might serve to stabilize this conformation during periods of reduced light or in darkness to ensure it is primed for rapid substrate turnover when photosynthesis recommences. Certainly, the rates of appearance and disappearance of the inhibitor are consistent with such a function.

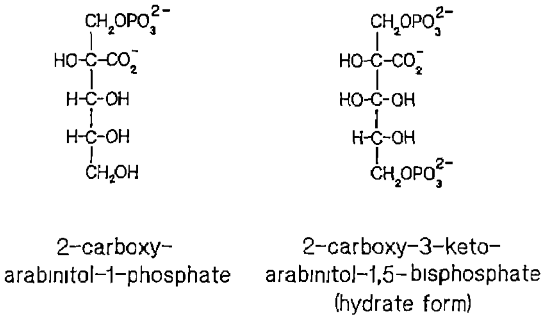

CH2OPO3 2-
HO-C-CO2 -
H-C-OH
H-C-OH
CH2OH
CH2OPO3 2-
HO-C-CO2 -
HO-C-OH
H-C-OH
CH2OPO3 2-
2-carboxy-arabinitol-1-phosphate
2-carboxy-3-keto-arabinitol-1,5-bisphosphate (hydrate form)

Fig. 3 The structure of the natural inhibitor 2-carboxyarabinitol-1-phosphate compared to a reaction intermediate of the carboxylase reaction of RuBPCase.

The metabolic precursor of the inhibitor has not yet been determined. The inhibitor must be a major component of dark-induced metabolism because the concentration of RuBPCase active sites in the stroma is at least 3 mM and therefore 2-carboxyarabinitol-1-phosphate must accumulate to these concentrations to cause significant inhibition. The inhibition of the enzyme at night has not been detected in all plants[3]. Until the biosynthetic enzymes that regulate the inhibitor concentration

Table 2 Chemical shifts* in p.p.m. relative to dioxane of the natural inhibitor, the dephosphorylated derivative and synthetic compounds

Compound	C-1 A	C-1 B	C-3	C-4	C-5 A	C-5 B
Natural inhibitor	3.73	3.71	3.82	3.52	3.64	3.40
2-Carboxyarabinitol-1-phosphate (open chain)	3.73	3.71	3.82	3.52	3.64	3.40
2-Carboxyarabinitol-1-phosphate (lactone)	3.82	3.72	4.19	4.24	3.80	3.58
Dephosphorylated inhibitor	3.59	3.53	3.70	3.51	3.63	3.40
2-Carboxyarabinitol	3.57	3.53	3.70	3.51	3.63	3.40
2-Carboxyribitol	(3.49–3.55)		(3.49–3.55)	3.37	(3.49–3.55)	
2-Carboxyxylitol	3.58	3.53	(3.54–3.58)	(3.54–3.58)	3.35	3.33
2-Carboxylyxitol	3.60	3.48	3.52	3.70	3.42	3.38

* Overlapping multiplets are shown as a range of p.p.m. values.

are identified, it will be unclear whether the absence of inhibition is due to rapid turnover or the lack of an available pool of precursors Alternatively, RuBPCase enzymes from different plant species may have different affinities for 2-carboxyarabinitol-1-phosphate, either intrinsically or because of competition with other endogenous phosphate esters Nevertheless, it is likely that this molecule is the natural inhibitor common to all plants that exhibit diurnal RuBPCase regulation In support of these observations, we have been informed that 2-carboxyarabinitol-1-phosphate has also been identified as the inhibitor bound to ribulose bisphosphate carboxylase in *Phaseolus vulgaris* at night[10]

We thank Roger Moffitt and the farm staff at Rothamsted for allowing unscheduled collection of leaves from potato plants and Dr N F Janes for discussion of the chemical evidence

Received 2 June, accepted 24 September 1986

1 Servaites, J C *Pl Physiol* 78, 839-843 (1985)
2 Vu, C V, Allen, L H Jr & Bowes, G *Pl Physiol* 73, 729-734 (1983)
3 Seeman, J R, Berry, J A, Freas, S M & Krump, M A *Proc natn Acad Sci USA* 82, 8024-8028 (1985)
4 Gutteridge, S & Keys, A J in *Topics on Photosynthesis* Vol 6 (eds Barber, J & Baker, N R) 259-285 (Elsevier, Amsterdam, 1985)
5 Gutteridge, S, Parry, M A J, Schmidt, C N G & Feeney, J *FEBS Lett* 170, 355-359 (1984)
6 Pierce, J, Barker, R & Tolbert, N E *Biochemistry* 19, 934-942 (1980)
7 Schloss, J V & Lorimer, G H *J biol Chem* 257, 4691-4694 (1982)
8 Schmidt, C N G *et al Photosynth Res* 5, 47-62 (1984)
9 Redgewell, R J *Analyt Biochem* 107, 44-50 (1980)
10 Berry, J A, Lorimer, G H, Pierce, J, Meeks, J & Freas, S M *Proc natn Acad Sci USA* (in the press)

Rearrangement and enhanced expression of c-*myc* in hepatocellular carcinoma of hepatitis virus infected woodchucks

Tarık Moroy*, Agnès Marchio*, Jeanne Etiemble*, Christian Trépot†, Pierre Tiollais* & Marie-Annick Buendia*

* Unité de Recombinaison et Expression Génétique (INSERM U163, CNRS US271), Institut Pasteur, 28 rue du Dr Roux,
75?24 Paris Cedex 15, France
† Unité de Recherches sur les Hépatites, INSERM U271,
UER Alexis Carrel, rue Guillaume Paradin, 69372 Lyon Cedex 3,
France

Hepatocellularcarcinoma (HCC) that occur in woodchucks chronically infected with woodchuck hepatitis virus (WHV) were screened for activation of cellular oncogenes. Enhanced expression and allelic alterations of the c-*myc* oncogene were found in three HCC out of nine. Variations in the size of the c-*myc* transcripts, ranging from 2.0 kilobases (kb) to 5 6 kb, as well as in the level of c-*myc* gene expression, 5–50-fold higher than in adjacent liver tissues, were observed among the three HCC. Rearrangements of the c-*myc* locus were either upstream of the gene or within the first intron. Cloning and sequencing of the break-point region from one of the three tumours showed that the c-*myc* gene was truncated and joined to a unique cellular sequence of unknown function. WHV DNA was not integrated near the c-*myc* coding exons, excluding a direct role of the virus in c-*myc* activation. The novel type of rearrangement and activation of the c-*myc* gene, reported here in liver tumours of hepatitis virus infected animals, appears strikingly similar to those resulting from chromosomal translocations in human Burkitt's lymphomas, acute B- and T-cell leukaemias and mouse plasmacytomas.

Total RNA recovered from HCC and adjacent liver tissues of woodchucks chronically infected with WHV was screened for expression of cellular oncogenes by RNA dot-blot hybridization and Northern blotting Similar low levels of transcription were observed in tumours and non-tumorous liver for most oncogenes tested (*abl*, *erb-A*, *erb-B*, *ets*, *fes*, *fms*, *myb*, *mos*, *raf*, Ha-*ras*, Ki-*ras*, N-*ras*, *sis* and *src*, not shown) The level of c-*myc* transcripts was markedly increased in three of nine HCC compared with the non-tumorous parts of the same livers (Fig 1) Polyadenylated RNA from the tumorous and non-tumorous parts of the livers of woodchucks W74, W93 and W64 was analysed by Northern blotting (Fig 1) using probes representa-

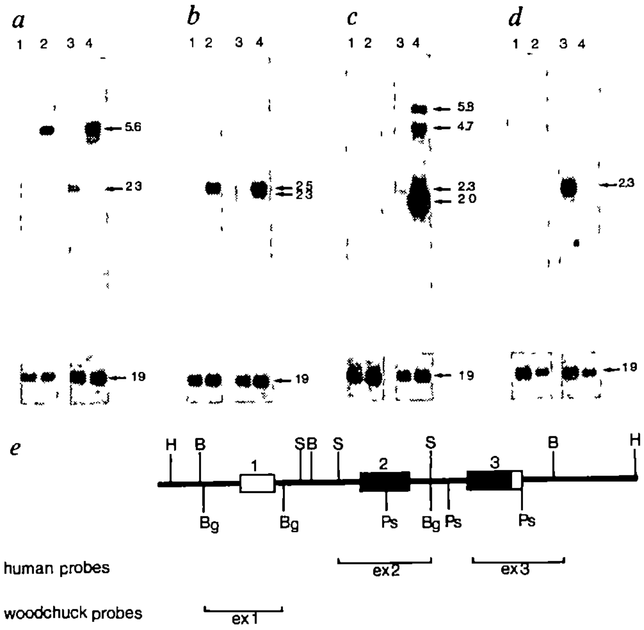


Fig 1 Northern blot analysis of c-*myc* mRNA from three woodchuck HCC and adjacent liver tissues Polyadenylated RNA from HCC (lanes 2, 4) and adjacent liver tissues (lanes 1, 3) of woodchucks (*a*, W74, *b*, W93, and *c*, W64), *d*, poly(A)$^+$ RNA from the human promyelocytic cell line HL60[24,25] (lanes 1, 3) and normal liver of woodchuck W20 (lanes 2, 4), Lanes 1 and 2 were probed with c-*myc* exon 1, lanes 3 and 4 were probed with c-*myc* exon 2 Restriction enzymes B, *Bam*HI, Bg, *Bgl*II, E, *Eco*RI, H, *Hin*dIII, Ps, *Pst*I and S, *Sac*I *e*, Restriction map of a molecularly cloned woodchuck c-*myc* gene

Methods *a*–*d*, Poly(A)$^+$ RNA was prepared as described elsewhere[21] Two identical samples of each RNA (5 μg for *a* and *d*, 10 μg for *b* and *c*) were electrophoresed in the same formaldehyde-agarose gel and blotted to nitrocellulose filters as previously described[26] The filters were prehybridized for 4 h at 42 °C in 50% formamide, 5 SSC, 50 mM NaH_2PO_4 *p*H 6 5, 1 × Denhardt's solution and 250 μg ml^{-1} sonicated, denatured salmon sperm DNA They were hybridized for 16–24 h with 10^6 c p m ml^{-1} of ^{32}P-labelled nick-translated probes (specific activity was routinely 2–5 × 10^8 c p m), after addition of 1/4 volume of 50% dextran sulphate The filters were washed twice in 0 1 SSC, 0 1% SDS at 60 °C for 30 min for hybridization with c-*myc* exon 1, or in 0 2 SSC, 0 1% SDS for hybridizations with c-*myc* exon 2, dried and exposed to Kodak XAR5 films with an intensifying screen at −70 °C for 48 h The filters were subsequently hybridized with the murine *β*-actin gene[27] to check for uniform loading (lower panels *a*–*d*) *e*, The positions of exons 1, 2 and 3 (boxes) were determined by hybridization with murine exon 1 and human exons 2 and 3 The boundaries of woodchuck *myc* exons were tentatively located by comparing conserved restriction sites in the woodchuck, human and murine c-*myc* genes, assuming similar organization of c-*myc* in the three species Only relevant restriction sites are shown Human c-*myc* exons 2 and 3 probes were a 1 5-kb *Sac*I fragment and a 1 6-kb *Cla*I fragment purified from a plasmid constructed in pKH47, a gift from Dr Stehelin The c-*myc* probe containing exon 1 was a 1 4-kb *Bgl*II fragment purified from a cloned woodchuck c-*myc* gene Hybridization signals obtained with a probe of murine c-*myc* exon 1 (a 0 4-kb *Bam*HI/*Sac*I fragment of pMc-myc 54 (ref 28)) were identical in position but very weak, indicating substantial divergence between murine and woodchuck c-*myc* exons 1, and therefore are not shown

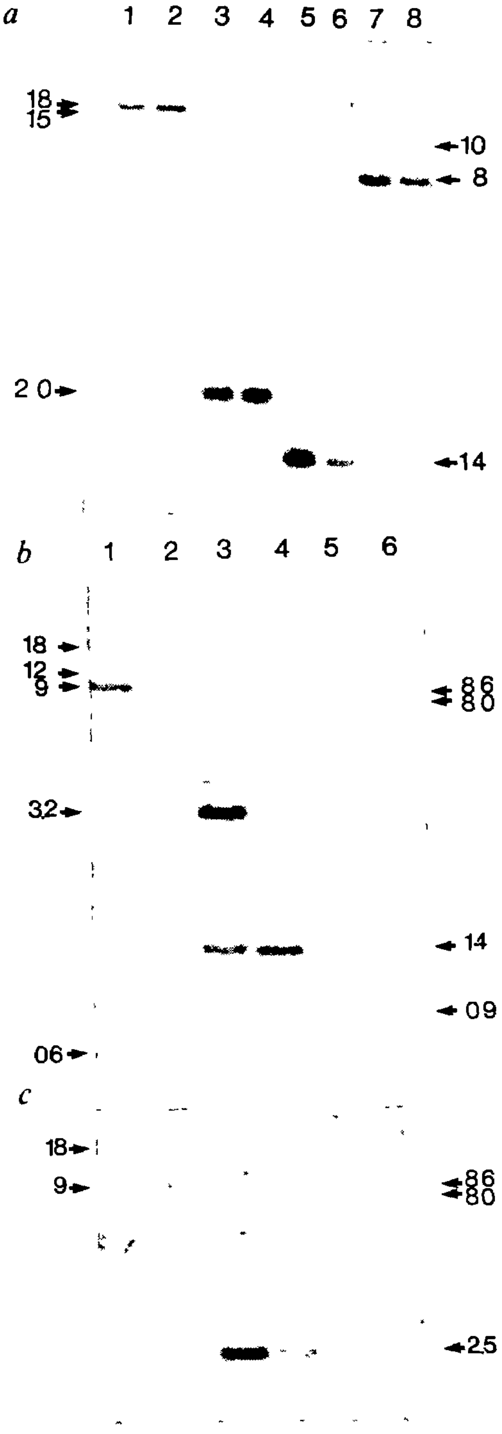


Fig 2 DNA rearrangement at the c-*myc* locus in W74 HCC and W93 HCC *a*, Hybridization of W74 HCC and W20 liver DNA with c-myc exon 1, *b*, hybridization of W93 tumour and non-tumour liver DNA with c-*myc* exon 1, *c*, subsequent hybridization of the same filter with c-*myc* exon 2 *a*, Cellular DNA (30 μg) from W74 HCC (lanes 1, 3, 5, 7) or W20 normal liver (lanes 2, 4, 6, 8) was digested with 1, 2, *Eco* RI, 3, 4, *Bam*HI, 5, 6, *Bgl*II, 7, 8, *Hin*dIII *b,c*, DNA from W93 tumour (lanes 1, 3, 5) and non-tumour (lanes 2, 4, 6 liver samples was digested with 1, 2, *Eco*RI, 3, 4, *Bgl*II, 5, 6, *Hin*dIII Southern blotting was performed as described[21] For probes and conditions of hybridization see Fig 1 legend The restriction pattern of DNA from the non-tumorous part of W74 liver was identical to that of W20 liver DNA and thus is not shown The sizes of the c-*myc*-containing fragments are given in kb DNA size markers were bacteriophage λ DNA digested with *Hin*dIII and double digested with *Eco*RI and *Hin*dIII The blots were exposed to X-ray films for 5 days The filters were reprobed with ^{32}P-labelled pBR322 DNA to ensure that all bands were *myc*-specific

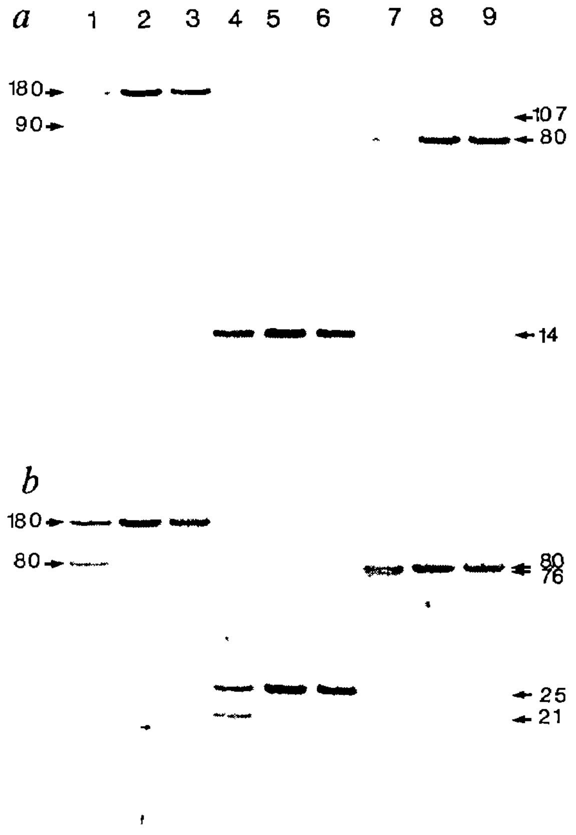


Fig. 3 DNA rearrangement at the c-*myc* locus in W64 HCC *a*, Southern blotting analysis of genomic DNA from W64 HCC (lanes 1, 4, 7), non-tumorous parts of W64 liver (lanes 2, 5, 8 and W20 normal liver (lanes 3, 6, 9) probed with *myc* exon 1 Lanes 1-3 are digestions with *Eco*R1, lanes 4-6 with *Bgl*II, and lanes 7-9 with *Hin*dIII (Figs 1 and 2 legends) *b*, The same filter was rehybridized with *myc* exon 2 Hybridization of the filter with a probe derived from pBR322 did not reveal any band Sizes in kb

tive of woodchuck c-*myc* exon 1 and human c-*myc* exons 2 and 3 The results obtained with *myc* exon 3 probe (not shown) were identical to those of *myc* exon 2 Densitometric analysis indicated that c-*myc* RNA levels were elevated at least 5-, 10- and 50-fold in W74, W93 and W64 HCC, respectively compared with the corresponding adjacent liver tissues from the same animal The level of c-*myc* transcription was below the threshold of detection in normal adult woodchuck liver The same probes revealed abundant 2 3-kb c-*myc* transcripts in the RNA of human HL60 cells (Fig 1*d*) However, hybridization of HL60 RNA with a c-*myc* exon 1 probe of woodchuck origin gave a considerably weaker signal than with the human c-*myc* exon 2 probe.

In contrast with the apparently normal 2 3-kb c-*myc* RNA found in the non-tumorous parts of the livers, the tumours produced c-*myc* transcripts of different sizes W74 HCC produced a major 5 6-kb c-*myc* RNA which annealed to all three c-*myc* exons (Fig 1*a*) The c-*myc* transcripts in W93 HCC also hybridized with all three c-*myc* exons (Fig 1*b*), they migrated as a heterogeneous population of 2 3-2 5 kb, slightly larger than normal This could reflect a shift in c-*myc* promoter usage from P2 to P1, as seen in Burkitt's lymphomas[1,2] and murine plasma cell tumours[3] W64 HCC had a major truncated 2 0-kb c-*myc* messenger RNA, and two minor mRNAs of 5 8 and 4 7 kb (Fig 1c) These RNAs annealed to human c-*myc* exons 2 and 3, but not to woodchuck c-*myc* exon 1 (Fig 1*c*), indicating that the normal transcriptional unit was disrupted in the tumour as has been reported for B-lymphoid tumours[4-7]

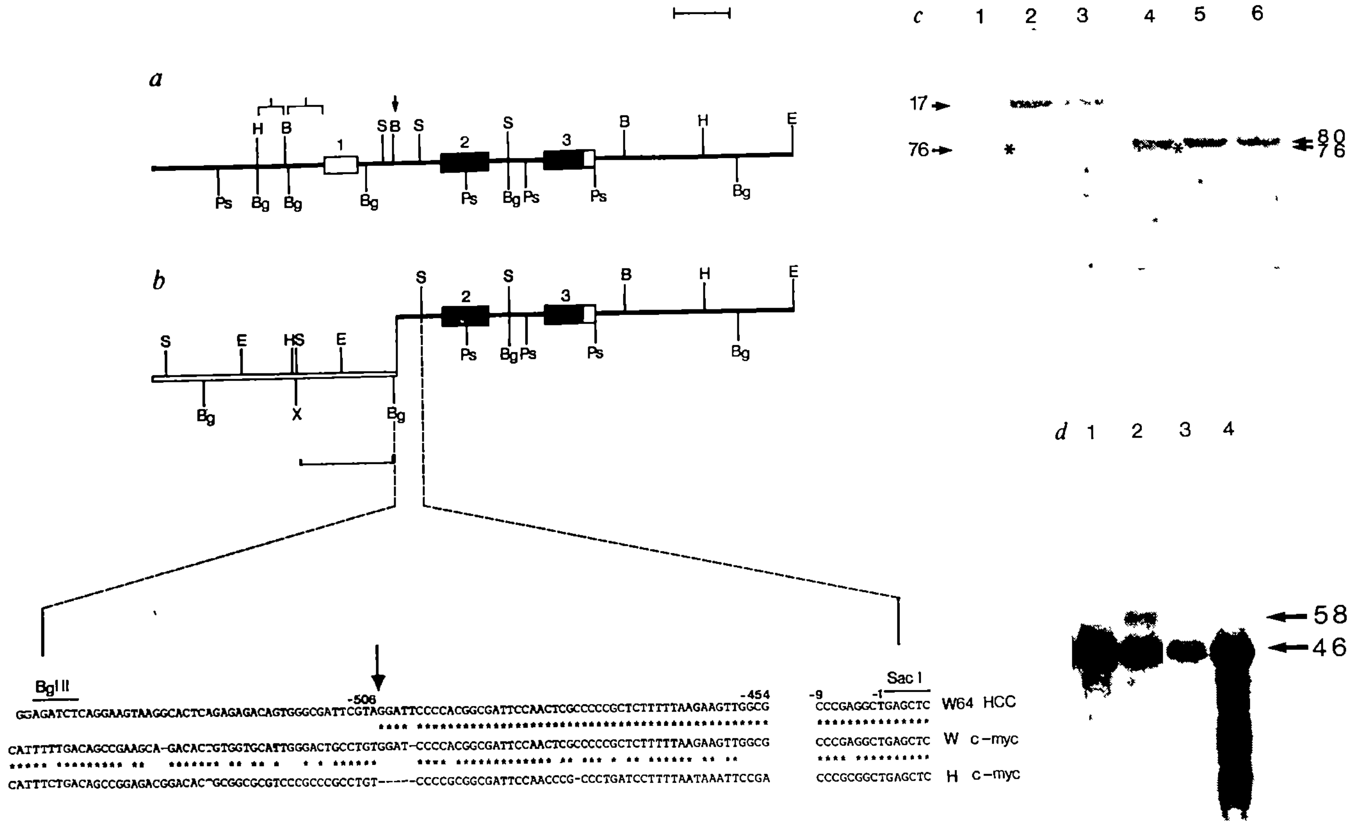

a
H B
1
SB S
2
S
3
B
H
E
Ps Bg Bg
Bg
Ps
Bg Ps
Ps
Bg
b
S
S
B
H
E
S
E
HS
E
Bg
X
Bg
Ps
Bg Ps
Ps
Bg
Bgl II
-506
-454
GGAGATCTCAGGAAGTAAGGCACTCAGAGAGACAGTGGGCGATTCGTAGGATTCCCCACGGCGATTCCAACTCGCCCCCGCTCTTTTTAAGAAGTTGGCG
CATTTTTGACAGCCGAAGCA-GACACTGTGGTGCATTGGGACTGCCTGTGGAT-CCCCACGGCGATTCCAACTCGCCCCCGCTCTTTTTAAGAAGTTGGCG
CATTTCTGACAGCCGGAGACGGACAC-GCGGCGCGTCCCGCCCGCCTGT-----CCCCGCGGCGATTCCAACCCG-CCCTGATCCTTTTAATAAATTCCGA
Sac I
-9
-1
CCCGAGGCTGAGCTC
CCCGAGGCTGAGCTC
CCCGCGGCTGAGCTC
W64 HCC
W c-myc
H c-myc
c
1
2
3
4
5
6
17
76
80
76
d
1
2
3
4
58
46

Fig 4 *a*, Schematic representation of the normal woodchuck c-*myc* allele Brackets, the recombination break points for the W74 HCC and W93 HCC as deduced from genomic blot studies, arrow, the precise breakpoint for W64 HCC (determined by molecular cloning and sequencing) The normal c-*myc* gene was mapped from restriction endonuclease digests of cloned DNA from W64 HCC as described (Fig 1 legend) Scale bar, 1 kb *b*, Mutant c-*myc* allele of W64 and sequence analysis of the break-point region Unshaded bar, incoming cellular sequences, small bar, probe s/bg, lower panel, enlargment of the region between *Bgl*II and *Sac*I sites The sequence of the mutant c-*myc* allele in the break-point area (W64 HCC) was aligned with the corresponding sequences of the normal c-*myc* intron 1 from woodchuck (W c-*myc*) and man (H c-*myc*) Asterisks, homologous nucleotides, arrow, breakpoint for W64 HCC *c*, Southern blotting analysis of *Hin*dIII-digested DNA from lanes 1, 4 W64 HCC, 2, 5, W64 non-tumorous part of liver, 3, 6, W20 normal liver Hybridization was performed as described (Fig 2 legend) Lanes 1-3 were probed with the 1 7 kb *Sac*I-*Bgl*II fragment located immediately upstream of the break point (small bar, Fig 4*b*), lanes 4-6 were probed with c-*myc* exon 2 Asterisks denote the tumour-specific fragment of 2 6-kb which also hybridizes with exon 2 (see Fig 3) *d*, Northern blotting analysis of total cellular RNA Lane 1, non-tumorous part and lane 2, tumorous part of W64 liver, lane 3, normal liver W101, lane 4, normal liver, W20 Hybridization to the s/bg probe was as described in Fig 1 legend

Methods Total cellular DNA from W64 HCC was partially digested with *Sau*3A 10-20-kb fragments were selected and ligated to the *Bam*HI-digested arms of the vector λ-L47 1 A library (2×10^6 phages) was screened with the human c-*myc* exons 2 and 3 (see Fig 1*e*), using the Benton and Davies *in situ* hybridization procedure[29] Eleven independent clones were obtained and analysed by restriction mapping and hybridization Five clones covered the entire normal c-*myc* gene as represented in Fig 4*a* Three clones retained c-*myc* exons 2 and 3 but contained non-*myc* sequences upstream of exon 2 The restriction map of the mutant c-*myc* allele was established from the DNA of these three clones (*b*), and is consistent with the data obtained by Southern blotting (Fig 3) The 2 3-kb *Sac*I fragment of the mutant allele, encompassing the break-point region was subcloned in a plasmid vector and further cloned into the M13 phage vector mp8, using the method of Deininger *et al*[30] The 0 7-kb *Sac*I fragment from the normal c-*myc* intron 1 was directly subcloned in M13 tg130 The sequencing reactions were performed as described by Sanger *et al*[31]

Cellular DNA from W74, W93 and W64 HCC was analysed for rearrangement or amplification of the c-*myc* locus Figure 2 compares Southern blots of genomic DNA from W74 and W93 HCC with W20 normal liver and with the non-tumorous part of W93 liver, respectively Additional c-*myc* bands were revealed in these HCC, indicating that a c-*myc* allele has been rearranged in the tumour cells The novel DNA fragments were not due to polymorphism in the length of restriction fragments, since identical normal c-*myc* bands were detected in DNA samples from the non-tumorous parts of livers W74 and W93 and from W20 (Fig 2) and 16 other woodchuck livers (not shown) Using woodchuck *myc* exon 1 as a probe, tumour-specific fragments were observed in *Eco*RI and *Hin*dIII digests of W74 HCC DNA, but not in *Bam*HI and *Bgl*II digests (Fig 2*a*) Hybridization with *myc* exons 2 and 3 revealed the same novel *Eco*RI and *Hin*dIII fragments and normal 2 5-kb and 4 1-kb *Bgl*II bands (not shown) Comparison of genomic blot analysis of W74 HCC DNA with the restriction map of a cloned normal woodchuck c-*myc* allele suggests a recombination event outside, and probably upstream of, the c-*myc* gene (Fig 4) In W93 HCC, novel 9-kb *Eco*RI and 8 6-kb *Hin*dIII fragments hybridized with *myc* exons 1 and 2 probes, and the faint 12-kb *Eco*RI and 0 9-kb *Hin*dIII fragments with *myc* exon 1 probe only (Fig 2*b*,*c*) Tumour-specific *Bgl*II bands were revealed only after hybridization with *myc* exon 1 This suggests a complex recombination event, next to the 5′ end of the c-*myc* gene

(Fig 4) Southern blot analysis of DNA isolated from different parts of W93 HCC revealed that rearrangements of the c-*myc* locus were restricted to one part of the tumour, this tumour was thus oligoclonal Enhanced expression of c-*myc* was found only in the tumour samples in which the c-*myc* gene was altered In contrast, W64 and W74 HCC appeared monoclonal after several blot experiments (results not shown)

W64 HCC also showed rearrangement at the c-*myc* locus (Fig 3) The probes for *myc* exons 1 and 2 annealed with the same normal *Eco*RI fragment, but with distinct tumour-specific bands This was also observed with *Hin*dIII digests The results obtained using a *myc* exon 3 probe (not shown) were identical to those for *myc* exon 2 Digestion with *Bgl*II revealed a novel band specific to *myc* exon 2 (Fig 3*b*) The c-*myc* DNA rearrangement in W64 HCC thus takes place within the first intron or the second exon The faint signals observed for tumour-specific fragments annealing to woodchuck *myc* exon 1 may indicate a partial loss of rearranged *myc* exon 1 We cloned the rearranged c-*myc* allele from a λ phage library of W64 HCC DNA, and located the recombination site in c-*myc* intron 1 by restriction mapping (Fig 4*b*) A 2 3-kb *Sac*I fragment, encompassing the joining region between c-*myc* and partner sequences, was subcloned and sequenced We compared this sequence with those of normal c-*myc* alleles from woodchuck (cloned from W64 HCC DNA) and man[8] Figure 4*b* (lower panel) shows that the breakpoint occurs 505 nucleotides upstream of the *Sac*I site nearest to the 5′ boundary of *myc* exon 2, which is found in both species Thus the mutant c-*myc* allele retains the cryptic promoters and initiation sites of intron 1, which are used for the transcription of truncated c-*myc* genes in human Burkitt's lymphoma[9], acute B-cell leukaemia[10], and murine plasmacytomas[5,7] In such cases, as in the woodchuck tumour, recombination *myc* exon 1 or intron 1 results in the production of novel RNA species that are shorter than the normal c-*myc* transcripts[11]

To improve our understanding of the mechanisms of c-*myc* activation in this tumour, we characterized the invading sequence next to the c-*myc* coding exons Hybridization studies of the cloned mutant c-*myc* allele from W64 HCC, using pWHV-1 as a probe[12], showed that there were no WHV sequences in the 5-kb regions flanking the 5′ and 3′ sides of the c-*myc* gene (not shown) A 1 7-kb *Sac*I-*Bgl*II fragment (denoted s/bg in Fig 4*b*) immediately next to the rearranged c-*myc* gene was used to probe Southern blots of W64 and W20 DNA, and revealed a unique 17-kb *Hin*dIII band in normal woodchuck liver DNA Another tumour-specific fragment of 7 6 kb in W64 HCC DNA annealed with both s/bg and *myc* exon 2 (Fig 4*c*) Further blot studies showed that the s/bg probe anneals to a unique sequence in the human genome (results not shown) The same probe revealed an abundant 4 6-kb transcript in W64 and normal W101 and W20 livers (Fig 4*d*) and other woodchuck livers (not shown) In W64 HCC, it also annealed to a 5 8-kb transcript migrating at the same position as the longer minor c-*myc* RNA which could therefore be a co-transcript of s/bg and c-*myc* exons 2 and 3 A computer-assisted search of the Los Alamos database failed to identify the s/bg sequence with any known cellular gene, including the immunoglobulin locus and the T-cell receptor gene

Activation of c-*myc* by different mechanisms has been implicated in a variety of haematopoietic malignancies and carcinomas[13,14] In some solid tumours, c-*myc* amplification[15,16] or rearrangement in a few cases[16] leads to an increased level of expression, but no major alteration of the c-*myc* locus has been reported in rodent liver tumours induced by chemical carcinogens[17] In B and T lymphoid tumours, c-*myc* expression is deregulated by a structural alteration of the gene, either chromosomal translocation[5,10] or insertional mutagenesis with non-acute oncogenic retroviruses[18,19] The DNA rearrangements and enhanced expression of c-*myc* in viral-associated HCC reported here seem quite similar to those found in B and T cell neoplasias It has been proposed that integration of hepatitis viruses in the cellular DNA of their natural hosts, observed almost invariably in liver tumour cells, is critical to HCC development[20] Indeed, Southern blot and cloning data revealed the presence of integrated WHV sequences in the three woodchuck HCC examined here (ref 21 and our unpublished results), but our data exclude c-*myc* activation by insertional mutagenesis in W64 HCC We have no evidence so far for viral insertion near c-*myc* in the other two tumours The recombination of a truncated c-*myc* gene with another cellular gene, described here for W64 HCC, is reminiscent of the reciprocal translocations of c-*myc* and immunoglobulin or T-cell receptor genes which occur in human Burkitt's lymphoma, B- and T-cell acute lymphoid leukaemia and mouse plasmacytoma Furthermore, it has recently been pointed out that in Burkitt's lymphoma and mouse plasmacytoma, c-*myc* alterations occur after a long pre-neoplastic period of chronic cell stimulation, provoked in some cases by Epstein-Barr or Abelson virus[22] A similar situation could result from chronic hepatitis virus infection in woodchuck and man, favouring chromosomal breaks[23] and finally leading to the development of HCC

We thank Dr S Aaronson, S Alonso, E Chang, T Curran, K Marcu, C Sherr and D Stehelin for providing *onc*-specific and actin clones, C Pichoud and C Jacquet for maintenance of woodchucks and surgical biopsies, M Hadchouel for histological examination of tumour tissue, J Robert-Lézènes for helpful discussions, J Davies for a critical reading of the manuscript and L M Da and A Cova for typing the manuscript T M is supported by the Commission of the European Community This work was supported by NIH grant CA 97300-02, the Economic European Community and the Association pour la Recherche contre le Cancer

Received 21 April, accepted 8 September 1986

1 Hollis, G F *et al.* *Nature* **307**, 752-755 (1984)
2 Taub, R *et al.* *Cell* **36**, 339-348 (1984)
3 Yang, J, Bauer, S R, Muchinski, J F & Marcu, K B *EMBO J* **4**, 1441-1447 (1985)
4 Shen Ong, G, Keath, E, Piccoli, S & Cole, M *Cell* **31**, 443-452 (1982)
5 Stanton, L W, Watt, R & Marcu, K B *Nature* **303**, 401-406 (1983)
6 Adams, J M, Gerondakis, S, Webb, E, Corcoran, L & Cory, S *Proc. natn. Acad. Sci. USA* **80**, 1982-1986 (1983)
7 Rushdi, A *et al.* *Science* **222**, 390-393 (1983)
8 Bernard, O, Cory, S, Gerondakis, S, Webb, E & Adams, J *EMBO J* **2**, 2375-2383 (1983)
9 Hayday, A C *et al* *Nature* **307**, 334-340 (1984)
10 Care, A *et al.* *EMBO J* **5**, 905-911 (1986)
11 Perry, R. P *Cell* **33**, 647-649 (1983)
12 Ogston, C, Jonak, G, Rogler, C, Astrin, S & Summers, J *Cell* **29**, 385-394 (1982)
13 Varmus, H E A. *Rev Genet* **18**, 553-612 (1984)
14 Weinberg, R A *Science* **230**, 770-776 (1985)
15 Yokota, J, Tsunetsugu-Yokota, Y, Battifora, H, Le Fevre, C & Cline, M J *Science* **231**, 261-265 (1986)
16 Escot, C *et al.* *Proc. natn. Acad Sci. USA.* **83**, 4834-4838 (1986)
17 Yaswen, P, Goyette, M, Shank, P R & Fausto, N *Molec. cell Biol.* **5**, 780-786 (1985)
18 Hayward, W S, Neel, B G & Astrin, S M *Nature* **290**, 475-480 (1981)
19 Corcoran, L M, Adams, J M, Dunn, A R. & Cory, S *Cell* **37**, 113-122 (1984)
20 Dejean, A, Bougueleret, L, Grzeschick, K. H & Tiollais, P *Nature* **322**, 70-72 (1986)
21 Moroy, T, Etiemble, J, Trepo, C, Tiollais, P & Buendia, M A *EMBO J* **4**, 1507-1514 (1985)
22 Klein, G & Klein, E *Nature* **315**, 190-195 (1985)
23 Rogler, C *et al.* *Science* **230**, 319-322 (1985)
24 Collins, S, Gallo, R & Gallagher, R Nature **270**, 347-349 (1977)
25 Collins, S & Groudine, M *Nature* **298**, 679-681 (1982)
26 Thomas, P *Proc. natn Acad. Sci. USA.* **77**, 5201-5205 (1980)
27 Alonso, S, Minty, A, Bourlet, Y & Buckingham, M *J molec. Evol.* **23**, 11-22 (1986)
28 Stanton, L W, Fahrlauder, P D, Tesser, P M & Marcu, K. B *Nature* **310**, 423-425 (1984)
29 Benton, W D & Davies, R. W *Science* **196**, 180-182 (1977)
30 Deininger, P L *Analyt Biochem.* **129**, 216-223 (1983)
31 Sanger, F, Nicklen, S & Coulson, A R *Proc. natn Acad Sci. USA* **74**, 5463-5467 (1977)

The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA

L. P. Eperon, J. P. Estibeiro & I. C. Eperon

Department of Biochemistry, University of Leicester,
Leicester LE1 7RH, UK

Alternative splicing of eukaryotic messenger RNA precursors is now known to be of widespread importance in generating multiple transcripts from a single gene. This phenomenon has emphasized the problem of the way in which splice sites are selected; recent studies have discussed the role of secondary structure[1] or affinity and spatial relationships[2] in this selection. Splice site sequences vary widely, although a loose consensus has been derived for the 9 bases around the 5′ splice site and for a longer region around the 3′ splice site[3]. Mutagenesis experiments have defined the sequences essential for a potential 5′ splice site[4,5,6], but, except for some experiments with the *E1a* gene of adenovirus[5,6], these experiments have not examined 5′ splice site sequences for features responsible for site preference where alternative splicing sites exist. Such tests require a choice of site: an appropriate reference site and a constant position at which test sites are introduced. We have begun a series of experiments designed to show whether splice site sequences can be ranked in a hierarchy of preferential use. Here we show that the archetypal consensus sequence is used efficiently, and characterize the cryptic sites of β-globin: sequences alone can explain why these sites are not normally used. We also show with the *E1a* gene of adenovirus, a simple example of alternative splicing, that one of the two 5′ splice sites used by this gene is intrinsically stronger We also demonstrate that tandem repeats and secondary structure influence the choice of sites *in vivo* We discuss the mechanism of splice site selection.

To define a hierarchy of splice site efficiency, we calibrated sequences by competition against a constant reference site An assay was devised involving choice between alternative splice sites contained in short oligonucleotides and the 5′ splice site of the second intervening sequence (IVS-2) of the rabbit β-globin gene (Fig 1*a*) The splicing junction to be assessed was always 25 nucleotides upstream of the authentic site The sequences tested (Table 1) vary at the 5′ splice junction in the last three nucleotides of the exon and first six of the intron, the region defined by the consensus sequence The 'cons' oligonucleotide is composed of the most frequently-used bases at each position in the consensus[3], 'glo' is a replica of the reference site, a necessary control for the effect of position, the 'cry' oligonucleotides contain sequences corresponding to the cryptic β-globin sites which are revealed only when the 'GT' at the beginning of

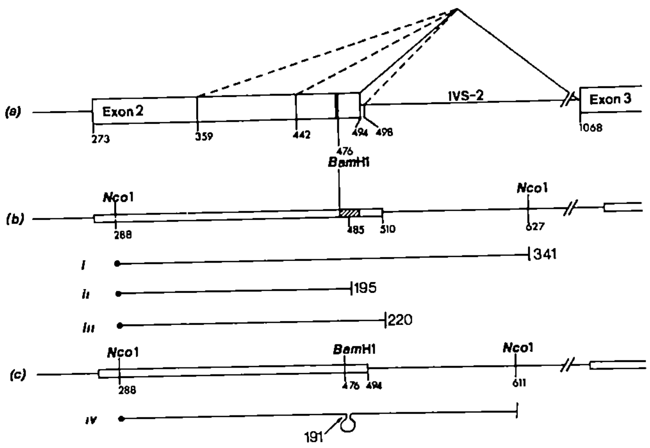


Fig. 1 *a*, Splicing of the second intervening sequence (IVS-2) of rabbit β-globin The boxes represent exons Solid lines between exons, authentic splicing pattern, dashed lines, splicing to cryptic sites when the 5′ splice site of IVS-2 is mutated *b*, explanation of the RNA mapping strategy, *c*, use of *in vitro* RNA corresponding to β-globin gene sequences *a*, The position of the *Bam*HI site into which 16-base oligonucleotide pairs (see Table 1) were cloned is shown The 'cons' and 'cry' paired oligonucleotides were first inserted into the *Bam*HI site of a *Bam*HI-*Bgl*II fragment of the β-globin gene which had been cloned into an M13 vector The *Bam*HI-*Eco*RI oligonucleotide-containing fragment was used to create the appropriate derivative of pβSV via three-way ligations The 'glo' oligonucleotides were inserted as a monomer, dimer (glo^2) and a tetramer (glo^4) directly into pβSV and the recombinants screened using the 'microprep' method This involved growing 96 small cultures in a microtitre plate and preparing the DNA by a modification of the alkaline lysis 'miniprep' method[11] The products were analysed by restriction enzyme digestion All the insertions into pβSV were sequenced using the dideoxy method[12] either directly on the double stranded plasmid or via recloning into M13 vectors *b*, An inserted oligonucleotide is shown by a shaded box, the *Bam*HI site is restored only at the 5′ end of the oligonucleotide (i), Undigested, full length probe, an *Nco*I restriction fragment, 3′ end-labelled by a two nucleotide extension at the staggered end using dCTP and [α-^{32}P]dATP and Klenow polymerase (indicated by a spot) This fragment was separated from other *Nco*I fragments on a 4% or 6% preparative sequencing gel containing 20% formamide (ii), The product after hybridization of the probe to an RNA transcript which has used the new 5′ splice site, and digestion with S$_1$ nuclease (iii), The digested probe when the reference site was used for splicing *c*, A *Pst*I-*Bgl*II fragment was cloned into pSP64, and RNA was trancribed *in vitro* from this (iv), Hybridization of a probe made from β-globin with oligonucleotide inserts to the RNA transcribed *in vitro* Digestion with S$_1$ nuclease would give a 191 nucleotide fragment The RNA made *in vitro* is a valuable size marker

Table 1 5′ Splice site efficiencies

Name	Sequence	Test site (%)	Reference site (%)	Details
glo	GATCCCAGG GTGAGTC	64	36	IVS-2 β-globin 5′ site
cons	GATCCCCAG GTAAGTC	100	0	From the optimal consensus[3]
cry1(498)	GATCCCTGA GTTTGGC	0	100	β-globin cryptics
cry2(442)	GATCCCTAA GCTGAGC	0	100	
cry3(359)	GATCCCAAG GTGAAGC	0	100	
glo^2		0	100	Dimeric IVS-2 β-globin, The 5′-most oligonucleotide is in the complementary orientation
glo^4		0	100	Tetrameric insert, the three 5′-most oligonucleotides are in the complementary orientation
E1a$_{973}$	GATCCCAGG GTGAGGC	0	100	Alternative E1a sites
E1a$_{1111}$	GATCCCACA GTAAGTC	63	37	
Ela$^2_{1111}$		100	0	2 correctly oriented oligonucleotides The 5′-most site is used
E1a$^{comp}_{1111}$		0	100	1 complementary sense oligonucleotide

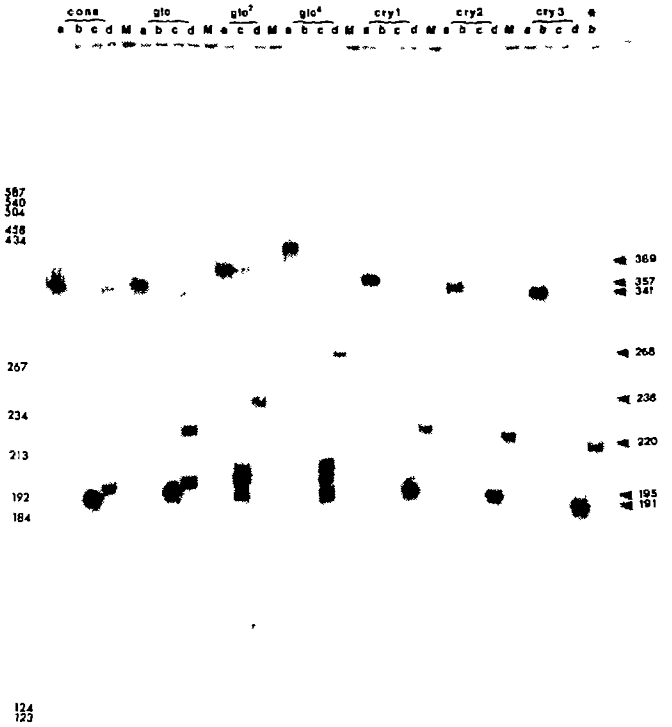


Fig 2 Determination of the splice sites chosen using S_1 nuclease, for β-globin genes containing the following test sequences the consensus (based on observed frequencies at each position[3] 43%C, 64%A, 73%G, 100%G, 100%T, 62%A, 68%A, 84%G, 63%T), β-globin IVS-2 (glo), IVS-2 dimer (glo^2), IVS-2 tetramer (glo^4) and the β-globin cryptic sites (cry 1, 2, 3) Total RNA was extracted from HeLa cells with LiCl$_2$-urea[13] 48 h after calcium phosphate transfection RNA was hybridized to the *Nco*I double-stranded probe at 49 °C for 2 h Standard conditions were used for S_1 nuclease digestions[14] Samples were separated on a sequencing gel[15] comprising 6% acrylamide, 20% formamide Size markers (M) were [^{32}P]-phosphorylated *Hae*III-cut pBR322 Their sizes are on the left hand side Numbers refer to base pairs For each construction, lane a shows the probe alone without S_1 digestion, lane b, S_1 nuclease digested probe, lane c, the relevant oligonucleotide containing probe hybridized to the *in vitro* β-globin transcript and lane d the result of hybridizing the probe to the expressed RNA followed by digestion with S_1 nuclease When the probe can form substantial secondary structure as in the case of glo^2 and glo^4 (lanes c) the pattern of S_1 nuclease digestion (multiple lengths) confirms that such structures are being formed Lane b for glo^2 (*) is at the far right hand side of the figure The autoradiographs were scanned with a laser densitometer after exposure to pre-flashed films (Kodak X-AR5) with an intensifying screen Lighter exposures than those shown were scanned, where the absorbance was proportional to the dose

IVS-2 is mutated[4], the *E1a* oligonucleotides ('E1a$_{973}$' and 'E1a$_{1111}$') are the two alternative 5′ splice sites of the adenovirus 2 *E1a* gene (at positions 973 and 1111) that are used to give two distinct mature transcripts Insertion of more than one oligonucleotide can be used to look at the effects of position and secondary structure

Oligonucleotides were cloned into the *Bam*HI site of a rabbit β-globin gene in pβ5′SV *Bgl*II[7] ('pβSV') and the products verified by sequencing The genes were introduced into HeLa cells and, after transient expression, the site of splicing was determined by S_1 mapping (Figs 1*b*, *c*) The proportion of splicing at the test site compared to the reference site was estimated by laser densitometry of autoradiographs

The results of the S_1 mapping experiments are shown in Figs 2 and 3, quantitative estimates are given in Table 1 A number of conclusions can be drawn

(1) The competitive assay for 5′ splice site sequences allows us to rank some sequences Splicing preferences can be shifted from the test site to the reference or vice versa depending on the sequence inserted at the test site The results with an oligonucleotide inserted in the complementary sense (E1a$^{\mathrm{comp}}_{1111}$) confirm that insertions of 16 nucleotides into the pre-mRNA structure do not disrupt splicing patterns This argues against a significant role for mRNA degradation induced by either the presence of extra upstream sequences when the reference splice site is used or the use of the new (test) site It is however possible that in some cases splicing to the reference site is followed by elimination of these transcripts, due to the presence of a splice signal at the test site (for example, 'cons' in Fig 2) Natural patterns of alternative splicing could be influenced likewise However, this would not impair the system's discrimination between sequences that promote splicing to the test site in the presence of the competing reference site and those that do not. However, if test site sequences have different susceptibilities to degradation, this could affect the comparisons among sequences that allow splicing to the test site (points (2), (3) and (6) below) Table 1 shows that the proportion of transcripts spliced to the reference site is inversely related to the degree of homology of a test site to the consensus sequence Thus degradation is probably insignificant or selective, in which case it would seem to reflect a further aspect of splicing activity at the test site In practice, the abundance of transcripts that have spliced to the reference site and carry natural splicing sequences at the test site suggests that splice site-specific degradation is not a major factor

A weaker reference site than that of β-globin may enable better discrimination among sites that are not used in the present arrangement

(2) The consensus sequence is used to the exclusion of the reference site, none of our other single inserts is as effective This site, which contains at each position the nucleotide found to be most abundant in natural sites[3], is so much 'stronger' than the others and small differences in sequence are so important (for example, cons and glo) that it is surprising all 5′ splice sites in genes that do not exhibit alternative splicing patterns do not conform exactly to the consensus The use of less efficient sequences would seem to imply unnecessary competition with exon sequences that by chance resemble naturally-occurring splicing signals However, such competition might facilitate the evolution of new proteins if acceptable signals arise The use of inefficient sequences might also alter the efficiency of gene expression One might expect that the time taken by splicing would be minimized to reduce degradation of nascent transcripts However, if site preference is related to the rate at which splicing occurs, then the frequent use of sub-optimal sequences suggests that there is little selective pressure for rapid splicing

(3) The consensus sequence and E1a$_{1111}$ differ only in the region that is usually within the exon, and these differences are clearly significant It has recently become common practice for 5′ splice site listings to include only the intron portion The region included in our oligonucleotides is the region for which conserved nucleotides have been detected, it would be surprising if sequences were conserved without functional importance We have not established whether sequences outside this region significantly influence splicing choices or efficiencies The exon sequences may constrain the evolution of protein sequence, or the protein sequence may dictate the splice site signals, this may not be important if sub-optimal sequences are acceptable in most genes (see point 2 above), but may influence alternative splicing patterns

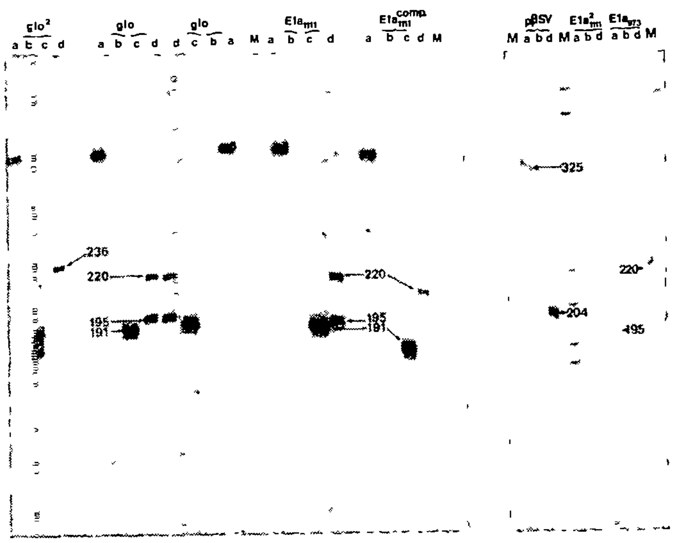


Fig 3 Determination of the splice sites chosen using S_1 nuclease, for β-globin genes containing the following test sequences rabbit β-globin IVS-2 5′ splice site (glo), adenovirus 2 *E1a* sites 973 and 1111 ($E1a_{973}$ and $E1a_{1111}$), a tandem dimer of $E1a_{1111}$ ($E1a^2_{1111}$) and a complementary insert ($E1a^{comp}_{1111}$), and no insert (p*βSV*) Lanes a, b, c, d and M are as described for Fig 2 Numbers refer to sizes in base pairs Samples were run on two 5% sequencing gels, containing 20% formamide, RNA was hybridized to the probe overnight at 49 °C and standard conditions were used for S_1 nuclease digestions[14] The two sets of lanes for glo represent independent isolates

(4) The sequences of the cryptic 5′ splice sites are not used, thus, the fact that these sites are not used in the β-globin gene (unless the normal IVS-2 5′ splice site is not available[4]) can be attributed to their sequence without invoking effects of secondary structure or position

(5) The results from constructions glo² and glo⁴ illustrate the importance of secondary structures Short inverted repeats can sequester a site and cause splicing to occur only at the reference This is a new way in which secondary structure can influence splicing, involving the site directly It has been shown[1] that an exon flanked by inverted repeats can be sequestered such that it is excised within the intron It is possible that the effect of secondary structure would be minimized if no choice was available, and we are testing this by inactivating the reference site in glo²

(6) Another factor which affects the choice of sites is that of relative position If we assume that the different flanking sequences of the two signals in the $E1a^2_{1111}$ tandem dimer have no bearing on their use, it seems that the upstream site is used preferentially and use of the reference site is drastically reduced This construction gave rise to a comparatively faint band on S_1 analysis (Fig 3), whether this is significant can be resolved only by experiments *in vitro* The significance of linear order may be influenced by the distance between two sites or by their environment, when a single β-globin sequence is inserted it is preferred to the reference β-globin site by a ratio of only 2 1

(7) The two *E1a* splice site sequences are used with very different efficiencies, even though both sites are used when intact *E1a* genes are expressed in HeLa cells Experiments in which β-globin restriction fragments were duplicated[2] showed that an upstream site is usually, but not always, used to the exclusion of a downstream site The results described in (6) emphasize the linear 5′ to 3′ order of alternative sites Thus, we anticipated that in the *E1a* gene a much more efficient sequence might be found at the downstream site in order to balance its unfavourable position This prediction was shown to be correct (Table 1, and unpublished results with sites intermediate in efficiency between glo and cryptic sites in our assay), but further experiments are required to test the importance of linear order and pre-mRNA structure

(8) The fact that splicing can be directed to the inserted oligonucleotide, away from a viable and efficient natural site, has mechanistic consequences If two 5′ splice site-binding complexes[8,9] bind simultaneously to the two sites, it is unlikely that the two sites could still be distinguished on the basis of their sequences in subsequent steps However, we have shown here that primary sequence alone is the basis of discrimination It is possible that both 5′ sites independently bind a complex and that the ratios of site usage are related to the average lifetime of bound complex We propose a model in which only one 5′ splice site complex can bind, perhaps limited by a requirement for the presence of a 3′ site complex (in agreement with some[8,10] but not all[9] *in vitro* data) The 5′ splice site complex can bind to either 5′ site, but the dissociation rates differ and the resulting pattern of usage then depends on the lifetime at each site The involvement of only one 5′ complex is rendered very likely by the proximity of our sites, experiments that increase the separation between the two 5′ splice sites will reveal whether the sequence discrimination and the model only apply to such close sites

We thank Mr J Keyte for synthesizing the oligonucleotides, and the members of the laboratory for critically reading the manuscript This work was supported by the MRC, J P E is a recipient of an SERC postgraduate training award

Received 27 May, accepted 24 September 1986

1 Solnick, D *Cell* **43**, 667-676 (1985)
2 Kuhne, T, Wieringa, B, Reiser, J & Weissman, C *EMBO J* **2**, 727-733 (1983)
3 Mount, S M *Nucleic Acids Res* **10**, 459-472 (1982)
4 Wieringa, B, Meyer, F, Reiser, J & Weissmann, C *Nature* **301**, 38-43 (1983)
5 Montell, C & Berk, A J *Nucleic Acids Res* **12**, 3821-3827 (1984)
6 Montell, C, Fisher, E F, Caruthers, M H & Berk, A J *Nature* **295**, 380-384 (1982)
7 Grosveld, G C, de Boer, E, Shewmaker, C K & Flavell, R A *Nature* **295**, 120-126 (1982)
8 Frendewey, D & Keller, W *Cell* **42**, 355-367 (1985)
9 Grabowski, P J, Seiler, S R & Sharp, P A *Cell* **42**, 345-353 (1985)
10 Perkins, K K, Furneaux, H M & Hurwitz, J *Proc natn Acad Sci USA* **83**, 887-891 (1986)
11 Estibeiro, J P (in preparation)
12 Sanger, F, Nicklen, S & Coulson, A R *Proc natn Acad Sci USA* **74**, 5463-5467 (1977)
13 Auffray, C & Rougeon, F *Eur J Biochem* **107**, 303-314 (1980)
14 Favaloro, J, Treisman, R & Kamen, R *Meth Enzym* **65**, 718-749 (1980)
15 Sanger, F & Coulson, A R *FEBS Lett* **87**, 107-110 (1978)

Errata

Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from Agrobacterium tumefaciens to plant cells

S. E Stachel, B Timmerman & P. Zambryski

Nature 322, 706-712

In this article, the title given on the contents page was incorrect The title is correct as above

A molecular link between the bats of New Zealand and South America

E. D Pierson, V. M. Sarich, J. M. Lowenstein, M J Daniel & W. E. Rainey

Nature 323, 60-63

In this letter, a line was omitted from the legend for Table 1, and *yumanensis* was misspelt The correct version reads

The taxa used in these comparisons represent all the families with which *Mystacina* has been associated *noctilio albiventris* (Noc) (Noctilionidae), *Mormoops megalophylla* (Mor) and *Pteronotus parnllii* (Moemoopidae), *Glossophaga soricina* (Glo) and *Phyllostomus discolor* (Phyllostomidae), *Antrozous pallidus* (Ant) and *Myotis yumanensis* (Vespertilionidae), *Tadarida brasiliensis* (Tad) (Molossidae), *Balantiopteryx plicata* (Bal) (Emballonuridae)

ANNOUNCING

Powdered Cell Culture

MEDIA

From HyClone Laboratories

- Low Endotoxin
- Biological Performance Tested On a Minimum of Three Cell Lines
- Components Carefully Selected and Tested
- Manufactured In Controlled Low Humidity Environment

For Further Information

***HyClone*®**

LABORATORIES, INC.

1725 So. State Hwy 89-91
Logan, Utah 84321
Toll Free 1-800-HYCLONE (492-5663)

Reader Service No.38

PLUS DE 200 NOUVEAUX PRODUITS

dont :

ANTICORPS MONOCLONAUX DIRIGES CONTRE LES ANTIGENES :

- de CELLULES HEMATOPOIETIQUES d'homme, de rat et de souris ;
- de CLASSES ET SOUS-CLASSE D'IMMUNOGLOBULINES d'homme, de rat et de souris ;
- de VIRUS, PARASITES ET BACTERIES ;
- d'HORMONES HUMAINES ET MARQUEURS TUMORAUX pour l'immunohistochimie.

SERACLONE : RENDEMENT DE FUSION SUPERIEUR A 90 %

- Lots de sérum fœtal spécialement sélectionné pour la fusion et le clonage cellulaire sur échantillon et réservations.

21, quai du Clos des Roses - 60200 Compiègne - France - Tél. : 44.86.22.75 - Télex : 145 171 F

Reader Service No.30

WHY PAY MORE FOR LESS?

CAMBRIDGE LIFE SCIENCES PRESENT THE 961 AND 962 MICROTITRE PLATE READERS.

A fresh approach by an innovative British Company.

Top quality instruments at about half the price of an average plate reader.

Benefits of the CLS 961 and 962 Models:

- Highly accurate and precise
- Interchangeable wavelengths
- Unique plate reading and transport
- Battery operated and rechargeable
- Compact, robust and portable
- Easy to use

The CLS 962 has these additional benefits:

- Auto blanking
- Sensitivity 0 - 2.999 Absorbance
- Data transmit facility with optional RS 232 output
- Optional rechargeable thermal printer

CLS

DIAGNOSTIC DIVISION

Cambridge Life Sciences plc
Cambridge Science Park Milton Road Cambridge
England CB4 4GN
Telephone: Cambridge (0223) 354144 Telex: 818713

Reader Service No.23

PRODUCT REVIEW

Immunoglobulin-derived drugs

Gary S. Hahn

Immunoglobulins may provide a novel source of drugs for the treatment and suppression of allergic reactions and other immunologically-mediated diseases.

RECENT research has suggested the possibility that small peptides derived from immunoglobulins (Igs) can serve as therapeutic drugs. All Igs secreted by plasma cells perform their effector and regulatory functions by a common pathway; by binding to Fc receptors (FcRs) in an isotype-specific manner. The precise FcR binding sites of Igs are unknown except for staphylococcal Protein A. Conventional dogma asserts that Ig FcR binding sites consist of large clusters of amino acids that require a specific three-dimensional configuration to bind to FcRs effectively. No small Ig-derived peptides with high affinities for FcRs have been reported to date. This historically led to the tacit assumption that small peptides from Igs would probably not be pharmacologically 'useful' since their affinity for FcRs would be too low.

Free-floaters

An alternative hypothesis, however, suggests that the Ig receptor binding sites consist of amino-acid clusters which also contain biologically active peptide segments. When released as free peptides, these small biologically active segments, or active site peptides (ASPs), have a low affinity for FcRs compared to intact Igs, and indeed may act as receptor agonists only after they are enzymatically released as free peptides (Fig. 1). According to this hypothesis, ASPs are released by proteolytic action while the Ig is bound to an FcR at the site of a developing immune response, and then act as short range immunoregulatory hormones.

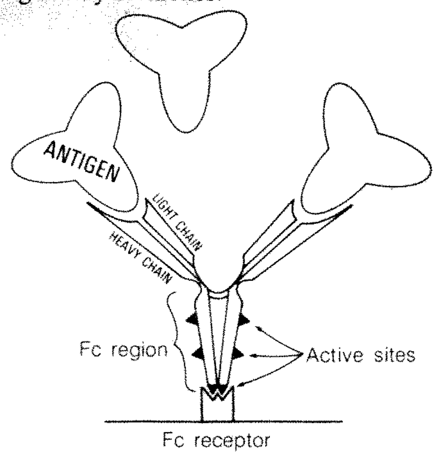


Fig. 1 Schematic of an antibody bound to two identical antigens and to an Fc receptor. The antibody's single Fc region contains multiple active site peptides (filled triangles) that bind to the Fc receptor and possibly to other immunoregulatory receptors.

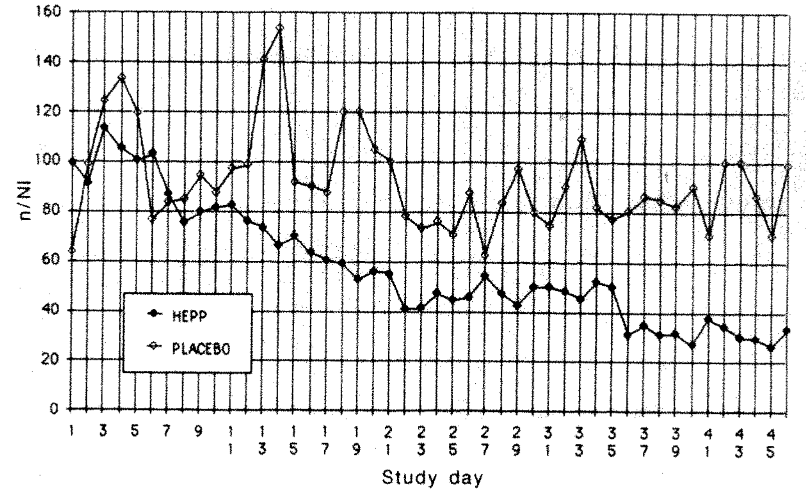


Fig. 2 Normalized stuffy nose intensity index (nIN1) during treatment with HEPP or placebo in 96 allergic rhintis patients.

A stumbling block may exist for the concept of Ig-derived ASP agonists though, as Igs are classically thought to trigger regulatory and effector functions only when two or more Igs cross-link multiple FcRs. This mechanism seems to exclude the possibility of FcR activation by small peptides, since the proposed peptides are too small to mimic FcR cross-linking. A critical point to consider in overcoming this conceptual hurdle is that the target receptors for ASPs may not necessarily be FcRs, but may instead be different immunoregulatory receptors. A low affinity FcR-binding peptide excluded as a potential ASP by conventional dogma is therefore an ideal candidate for an ASP according to these new concepts.

Small Ig-derived ASPs could have great potential as therapeutic drugs because, unlike Igs, ASPs are likely to be nonallergenic and inexpensive to manufacture. Most importantly, ASPs would probably be rapidly degraded to amino acids and would thus lack toxicity due to long-lasting metabolites.

The first ASP derived from any Ig was the tetrapeptide tuftsin isolated from enzymatic digests of human IgG. Tuftsin was shown to stimulate phagocytosis and other protective immune responses. As monomeric IgG does not stimulate phagocytosis, it was suggested that tuftsin must first be enzymatically released before expressing activity[1].

Subsequent studies have demonstrated the existence of other Ig-derived peptides whose immunoregulatory activities appear to be unrelated to traditional Fc/FcR interactions. The tetrapeptide rigin, and an unnamed decapeptide containing the rigin sequence, stimulate phagocytosis like tuftsin[2]. In addition, a 23-amino acid peptide enzymatically released from human IgG, containing both the rigin and the decapeptide sequences, was reported to stimulate T cell-dependent antibody synthesis with suboptimal concentrations of antigen[2]. A nonapeptide from the C terminus of human IgE (igercin) is also claimed to have kinin-like effects after enzymatic release, whereas an enzymatically released pentapeptide from IgM (IPM-5), located at a site in IgM spatially related to the rigin site in IgG, is claimed to exhibit thymic hormone-like activity[3].

Several ASPs have demonstrated therapeutic activity *in vivo*. Tuftsin given to mice with neoplastic tumours has been shown to increase the survival rate, and the 23-amino acid human IgG peptide mentioned above was reported to augment natural killer cell activity in mice. A short peptide from human IgG (73/He.2.80) substantially prevented experimental allergic encephalomyelitis in mice, as judged by neurologic deficit scores and by the reduction of central nervous system lesions, and is expected to be tested in human clinical trials on patients with multiple sclerosis[4].

Only one Ig-derived peptide, human IgE pentapeptide (HEPP), has been shown to have therapeutic activity in humans. HEPP's ability to inhibit IgE-mediated inflammation in humans was initially suggested to be due to competitive antagonism for the IgE FcR[5]. This concept was later proven untenable because of HEPP's low affinity for IgE FcRs. Figure 2 shows the effect of HEPP on 'stuffy nose' during a seven-week, double-blind placebo-controlled trial in 96 allergic rhinitis patients, in which the subjects recorded daily symptoms for one week prior to treatment and for six weeks during treatment (days 8–46)[6]. As an Ig-derived ASP, HEPP's activity is unexpected based on traditional concepts of Ig/FcR interactions.

The potential identity of some of the receptors for which Ig-ASP's may act as agonists is suggested by tantalizing sequence similarities of certain ASPs and other biologically active molecules. IPM-5, for example, is identical except for one amino acid to thymopentin, an ASP of the thymic hormone thymopoietin[3]. The first three amino acids of IPM-5 (Arg-Pro-Asp) are nearly the same as the binding site of fibronectin and its molecular relatives vitronectin, laminin and fibrinogen (Arg-Gly-Asp)[7]. The IgE-derived igercin resembles the inflammatory peptide bradykinin[3]. IgG-derived tuftsin resembles the N terminus of Substance P (SP-4) and has been shown to bind to Substance P receptors, and when injected intracerebroventricularly, tuftsin induces analgesia (Substance P antagonism). SP-4 has similarly been demonstrated to have tuftsin-like phagocytosis-stimulating activity[1].

New horizons

While certainly speculative, these curious resemblances suggest the possibility that Igs may have various releasable signals coded within their structures that act as short range hormones upon non-FcR systems. This greatly challenges the conventional concept of Ig structure and function, and demands explanations for the observed biological activities of ASPs that seem unrelated to known Fc/FcR interactions. The exciting possibility that Igs may serve as design templates for novel classes of immunoregulators may serve to redirect the thinking of molecular immunologists whose initial research may have revealed only a small fraction of the potential of Ig-derived peptides to serve as therapeutic drugs. □

Gary S. Hahn is an assistant clinical professor in the Immunology and Allergy Division of the Department of Pediatrics at the University of California, San Diego School of Medicine, and is vice president for research and development at Immunetech Pharmaceuticals, 11045 Roselle Street, San Diego, California 92121, USA.

1. Najjar, V.A. & Fridkin, M. *Ann. N.Y. Acad. Sci.* **419,** 1-272 (1983).
2. Hahn, G.S. in *Protein Abnormalities* Vol. 1 (ed. Ritzman, S.E.) 193–304 (Liss, New York, 1982).
3. Chipens, G. *Surv. immun. Res.* **4,** 220–229 (1985).
4. Armstrong, R.M., Bray, R.S. & Amstrong, D.L. presentation to the Canadian Neurologic Society International Meeting, June 1985.
5. Homburger, R.N. *Science* **189,**. 389–390 (1975).
6. Dennis, S. *et al. J. Allergy clin. Immun.* **77,** 213 (1986).
7. Leptin, M. *Nature* **321,** 728 (1986).

Immunotech ideas

More and more monoclonals are on the market, plus the equipment to produce them.

Becton Dickinson has introduced **simultaneous two-colour immunofluorescence** with the new Simultest line of monoclonal reagents (Reader Service No. 100). Simultest makes use of a combination of monoclonal antibodies conjugated with fluorescent dyes that are excited at the same wavelength, so that multiple tests can be run with one assay. Intended for research applications only, Becton Dickinson's reagents are currently available for identification of T cells, B cells and certain other lymphocyte populations. The cost for the Simultest T and B Cell Test is $285 (US), and it contains enough reagent for 50 runs.

Simultest makes cells glow

Culturing hybridomas for monoclonal antibody production requires a **serum-free medium**. ICN Biochemicals produces Hybri-Clone SF specifically for this application (Reader Service No. 101). Provided in liquid or lyophilized form, Hybri-Clone SF is also suited for growth of primary or established cell lines. A 100 ml bottle sells for $17.50 (US).

Serum-free for purer cultures.

A Brandel **multi-probe cell harvester,** suitable for receptor binding assays, is now available from Semat Technical, Ltd. (Reader Service No. 102). Models equipped with 12, 24, and 48 wells are available, that Semat says perform cell harvesting from most sizes of test tubes within 30 seconds. Options include a hot-cold valve to separate radioactive from nonradioactive waste, a wash timer, and a collection box allowing the individual collection of filtrate. The basic 12-well M-12R model has a price tag of £2,395 (UK).

To ensure the **recovery and concentration of monoclonal antibodies** and other recombinant proteins in their biologically active forms, Sterogene has developed its ActiSept Immunoaffinity Purification System (Reader Service No. 103). The complete system includes an activated resin immunoadsorbent with a flow rate of 170 ml cm^{-2} h^{-1} (for Actigel A-Superflow) that immobilizes antigens and antibodies at physiological *p*H, and a nondenaturing eluent that Sterogene claims allows recovery of active proteins with a 98 per cent yield. The basic kit of 100 ml of Actigel A or Actigel A-Superflow resin plus 1 litre of elution medium costs $270 (US). Immunogels containing anti-mouse and anti-rabbit antibodies, as well as Protein A-Actigels for the isolation of IgG, are also available.

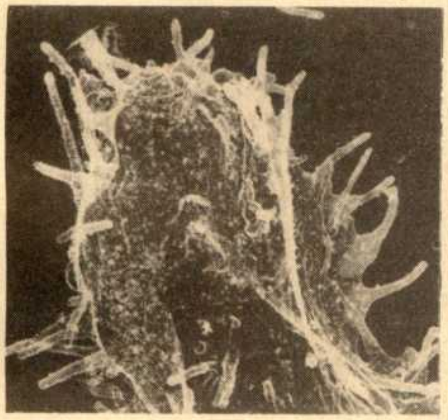
More monoclonals

For ease in identifying cell origins during cancer and other cell culture research, Du Pont Ltd offers its Cytolite line of **monoclonal antibodies specific to cytoskeletal proteins** (Reader Service No. 104). Du Pont's monoclonals are available for vimentin, actin, tubulin and keratin, allowing experimenters to assign major cell type origins. The monoclonals typically cost £84 (UK) for sufficient antibody for 50 assays, and each order is accompanied by instructions outlining the exprimental procedures for rhodamine or fluorescin detection.

Customized antibodies are available from BAbCO (Berkeley Antibody Company) (Reader Service No. 105). Its free 16-page guide outlines an extensive range of polyclonal and monoclonal antibody services. Ancillary immunological services are offered, including hapten-carrier conjugation, ELISA, and pharmacokinetic studies. BAbCO's services are keyed to researchers on any size budget, and preparations are made directly from client antigens.

MicroGeneSys, Inc. offers up-to-the-minute **products for AIDS research** purposes (Reader Service No. 106). Its MGStandard HIV gp160 protein is isolated from the expression of the virus envelope protein *env* in insect cells, and is thus free of contaminating mammalian cell proteins. The protein is complete except for a small deletion of approximately 100 amino acids at the C terminus, introduced for stabilization, that does not affect the two hydrophobic domains present near the C terminus of the gp41 portion. MGStandard HIV gp160 is supplied as a tested 50% pure solution in a sterile aqueous buffer at a concentration of 100 mg ml^{-1}, and a 25 mg sample costs $385 (US). For convenience, the AIDS envelope protein precursor is also available in prepared 96-well plates for ELISA, and in impregnated nitrocellulose strips (3–4 mm × 10 mm) for Western blot analysis.

Known for radiochemicals as well as for radioligands for use in receptor detection, Amersham International has recently expanded its product line to include a range of **monoclonal antibodies** (Reader Service No. 107). Amersham's new receptor monoclonals are specific for epidermal growth factor receptor, interleukin-2 receptor, and transferrin receptor. To accompany the antibodies, Amersham also sells detection systems using biotin-streptavidin, streptavidin–gold, and DAB silver enhancement.

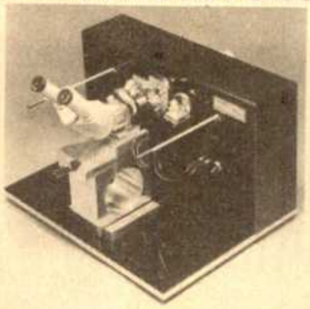


Transferrin receptors on human laryngeal carcinoma cells visualized with streptavidin-gold.

Flow Laboratories is now distributing Hybritech's expanded range of **monoclonal antibodies,** and Hybritech has recently announced three additions for research applications (Reader Service No. 108). Monoclonals for T-cell surface antigens include antibodies for interleukin-2 receptor for the detection and enumeration of activated T cells, andd those for suppressor/cytotoxic and TA 5.9 T-cell antigens, useful for lymphocyte phenotyping as well as T-cell activation. Immunohistochemical detection of human, monkey and porcine chromogranin, indicating the presence of endocrine-derived cells, is

ADVERTISEMENTS

SINGER INSTRUMENTS
MICRO-TOOL MAKING?
then you need the
SINGER MICROFORGE



Write to: Singer Instrument Co Ltd, Treborough Lodge, Roadwater, Watchet, Somerset TA23 0QL, England. Tel: (0984) 40226. Telex. 337492 Comcab G.

Reader Service No.50

Activity Monitor
Track that rat, insect, fish, limb
Tracking unit accepts input from video camera or recorder, and outputs the location of a high contrast object in the TV image. Interfaces for BBC, Apple, IBM-PC.

HVS IMAGE ANALYSING LTD,
22 Cromwell Road, Kingston,
KT2 6RE, England.
Tel: 01-546 9090

- NEW PRODUCT -
Cell-length Monitor

Reader Service No.13

Medwire Div.
stocks a wide variety of
TEFLON COATED*
Platinum alloys, stainless steel, silver, multistrand and single strand available for prompt shipment. Write for descriptive brochure and price lists.

Medwire Div.
Affiliate of Sigmund Cohn Corp.
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300
*Reg. T.M. of E.I. Dupont

Reader Service No.49

WESTERN BLOTS IN MINUTES
Our new GENIE electrophoretic blotter incorporates space-age plate electrodes for a uniform, high-intensity field. It is easy to load, and uses less than a liter of buffer, yet can blot gels up to 15 by 17 cm. in as little as ten minutes.
only $190
Idea Scientific Company
P.O. Box 2078
Corvallis, Oregon 97339
(503) 758-0999
ASCB Booth No. 326
Reader Service No.53

MONOCLONAL
ANTIBODIES

CELL SURFACE ANTIGENS
human T-cell T6 (CD1), clone 66-ILC7
human T-cell T3 (CD3), clone RIV9
human T-cell T3 (CD3), clone SPV-T3b
human T-cell T4 (CD4), clone RIV6
human T-cell T1 (CD5), clone 6.12.20.1 and clone 6.4.15.1
human T-cell differentiation antigen (CD7), clone WT1
human T-cell T8 (CD8), clone SPV-T8
human T-cell T8 (CD8), clone FK18
human T-cell T9/Transferrin receptor, clone 66-IIG10
human HNK-1, clone 6.13.19.1
human Monocyte/Granulocyte antigen (CD11), clone FK24
human Granulocytes (CDw15), clone FK32
HLA-D, clone 7.510.1
HLA-DQw1 related determinant, clone II B3
HLA-DRw52 related determinant, clone 7.3.19.1
human T-cell receptor, clone WT31

HORMONES
human Chorionic Gonadotropin, (hCG), clone 2092
human Follicle Stimulating Hormone, (FSH), clone 2093

INTERMEDIATE FILAMENTS
human Desmin, clone 33
human Glia Fibrillary Acidic Protein, clone 6F2
human Keratin, clone 80
human Neurofilament, clone 2F11
human Vimentin, clone 9

CYTOSKELETON ANTIGENS
actin, (rabbit muscle), clone SA1C1
tubulin, (rat), clone 1G9

PROTEINS
human Kappa, clone 2B7 and 3B10
human Lambda, clone 48
human Thyroglobulin, clone 14/14
human IgM, clone B8904 E7

HISTOLOGICAL ANTIGENS
human Endothelium , clone PAL-E
human Endothelium, clone EN4

MELANOMA ANTIGENS
Determinant discriminating between nevocellular nevi and malignant melanomas, clone PAL-M1
Determinant discriminating between nevocellular nevi and malignant melanomas, clone PAL-M2

GROWTH FACTORS
Epidermal growth factor (mouse), clone E5
Epidermal growth factor (mouse), clone F5

NEW MONOCLONAL ANTIBODIES
NEW PRODUCTS
NKI/C3 for diagnosis of Amelanotic Melanoma
For immuno histochemistry: PROSTATE SPECIFIC ANTIGEN
* Polyclonal antibodies against
SEROTONINE
DOPAMINE
Other antisera against neurotransmitters expected soon.
ENZYME LABELLED ANTIGENS.
① Toxoplasma antigen-HRP code TOX 8610
② Cytomegalovirus nuclear antigen-HRP code CMV 8620
③ Varicella-zoster virus antigen-HRP code VZV 8630

SANBIO bv
Sanbio bv - biological products
Heinsbergenstraat 50, 5402 EG UDEN
P.O. Box 540, 5400 AM UDEN, The Netherlands
Phone: +31 4132 - 51115. Telex: 74827 SANBO NL

Reader Service No.32

possible with Hybritech's monoclonal LK2H10. LK2H10 is also useful in Western blotting and purification of chromogranin from human adrenal glands. Monoclonal antibody to group B streptococci (Type III) is also available. Hybritech say their monoclonals are suitable for flow cytometric analysis as well as fluorescence staining, and all products are lyophilized and stable for at least one year. Prices range from £98 (UK) for a 0.2 mg sample of antibody for cell surface antigens, to £198 (UK) for 0.5 mg of antibody for tumour antigens.

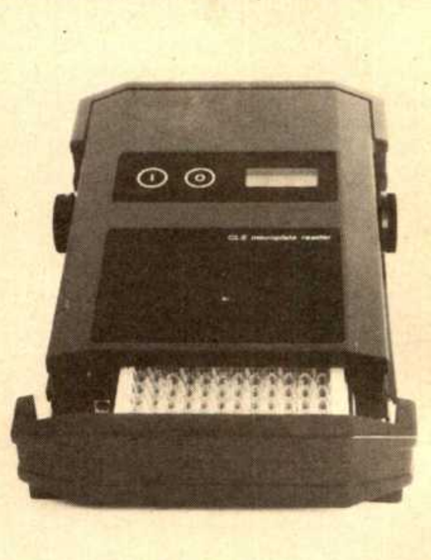
Microplate machines

A microprocessor-controlled Fluorescence Plate Reader is available from Perkin-Elmer as an accessory to be used with its LS-3B or LS-5B fluorescence spectrometers (Reader Service No. 109). Perkin-Elmer's suggested applications for its scanner include monitoring fluorescence immunoassays and monoclonal antibody screening. The £950 (UK) plate scanner operates by using fibre optics to transmit excitation and emission energy to and from the fluorescence spectrometer, as excitation energy absorbed by the sample produces fluorescence that is detected by the emission fibres in the light pipe head.

Microplate washing and aspirating can be tedious and time-consuming. Dynatech's Ultrawash II automates the process, and allows the operator to pre-select dispense volumes, soak times between 0 and 90 s, rates of aspiration and number of wash cycles (Reader Service No. 110). The £2,412 (UK) machine has 96 channels, and incorporates alternative dispense/aspirate heads to allow the selection of single or multiple rows as desired.

Dynatech's Ultrawash II is designed for those who hate to wash plates.

Cambridge Life Sciences is making **microtitre plate readers** affordable, claiming that its reader is half as expensive as the average instrument, with top quality (Reader Service No. 111). The fully portable CLS 961 is provided with a carrying case and costs £895 (UK). The CLS 961 reads 96-well ELISA and other microwell assays, and offers a choice of wavelengths of 405, 450, 492, 540, 615 and 690 nm. The more expensive CLS 962 has greater sensitivity and a wider absorbance range, and can be connected to a computer or printer.



Carry the CLS 961 anywhere.

ADVERTISEMENT

Fitting Solutions

. . .To your reagent needs with Becton Dickinson Monoclonal Antibodies. The largest available range of antibody formats for research and diagnostic work provides the solution, whatever the pattern of your workload.

Another example of innovative products resulting from Becton Dickinson's commitment to research and development for the clinical laboratory, the BD Monoclonal Antibody range is the widest and of the highest quality available for research and diagnostic work on leukocytes in tissue and blood. Supplied ready to use in solution, with a minimum shelf life of six months, and two years on purified immunoglobulins at 2-8°C, they combine guaranteed reliability and quality with cost-effectiveness. Convenient to use and with the best possible supporting user literature, BD MABS are used by leading scientists in the fields of AIDS, leukaemia, cancer, multiple sclerosis and transplant research.

What more fitting solution to your Monoclonal Antibody needs?

- Anti -IL2 Receptor : Activated T. cells and HTLV infected cells.
- Anti -HPCA–1 : Human progenitor cell associated antigen, reacts with acute leukaemia specimens and unipotent and multipotent progenitor cells.
- Anti - Cytokeratin (CAM 5.2) : reacts with adenocarcinomas and mesotheliomas
- Anti - Calla : human common acute lymphoblastic leukaemia antigen.

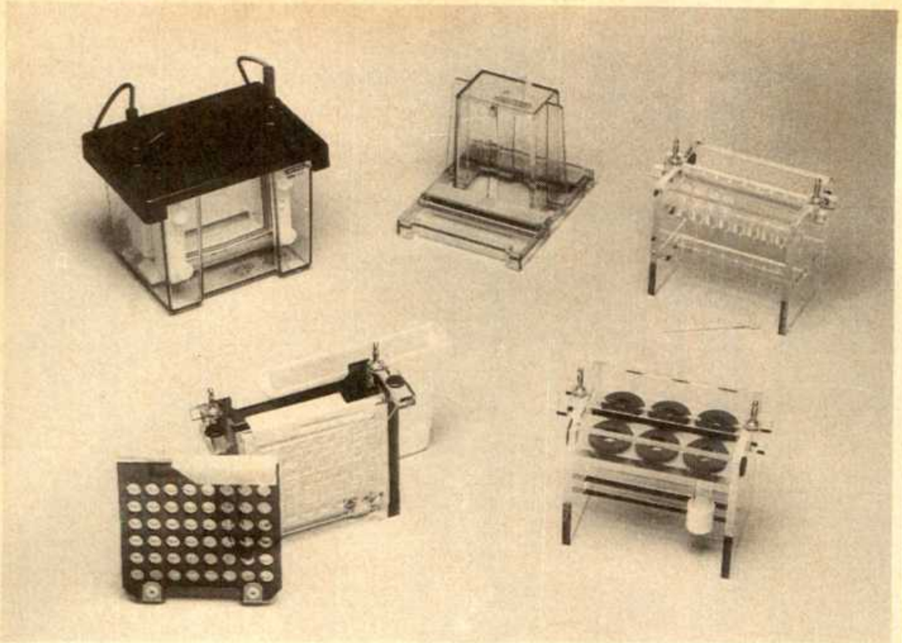
Further information available from:
Laboratory Impex Limited, Impex House,
Lion Road, Twickenham, Middlesex TW1 4JF
Tel: 01-891 4881 Telex: 8951670 Labimp G

Laboratory Impex Ireland Ltd, 3A Kylemore Park North,
Ballyfermot, Dublin 10. Tel: Dublin 262332

BECTON DICKINSON

Laboratory Impex Limited

Reader Service No.6

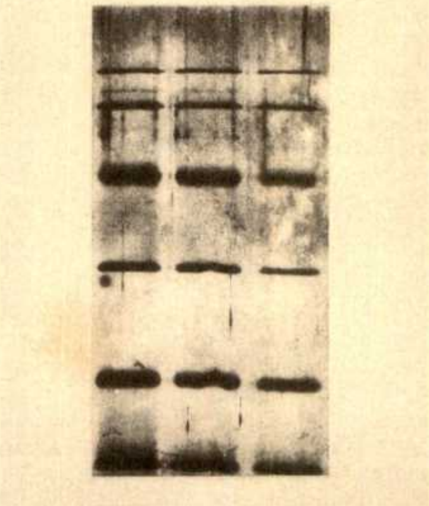


Bio-Rad's array of modular electrophoresis products.

Collecting, detecting, and dispensing

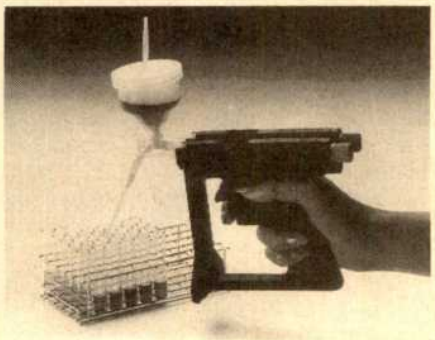
Bio-Rad has developed three **interchangeable electrophoresis modules** to expand its Mini Protean II electrophoresis cell into a complete system (Reader Service No. 112). The insertion of the Mini Trans-Blot module after the initial electrophoretic separation is completed allows electrophoretic transfers to be performed within an hour, according to Bio-Rad. By switching the module again, to the Mini Tube Cell, first dimension IEF tube gels for 2-D applications are produced. With another change of modules, the Electro-eluter recovers proteins and nucleic acids from the gel. Each module may be purchased separately, or the entire system, including the Mini Protean II basic cell, is available for £1,075.50 (UK).

Bio-Rad also offers a **colloidal gold protein detection kit** for the detection of proteins bound to nitrocellulose membranes (Reader Service No. 113). The Enhanced Colloidal Gold Total Protein Detection Kit is sensitive to as little as 400 picograms of bound protein, and includes the Colloidal Gold Protein Stain, Tris, Tween-20, a Gold Enhancement Kit, and an instruction manual. The Gold Enhancement Kit deposits a layer of silver over all membrane-bound gold molecules, turning the band from red to black, and thereby increasing its visibility. The total assay time ranges from two hours to overnight. The kit has a sensitivity equivalent to silver staining of polyacrylamide gels, according to Bio-Rad, and provides protein detection blots identical in size to the immunostained membrane. Bio-Rad claims that its kit also eliminates the shrinking problem associated with anionic dyes.



Enhanced detection with gold and silver.

Manual **sample dispensing and dilution** is the purpose of Labsystems' Finnpipette Diluter (Reader Service No. 114). The hand-held gun has one chamber for sample and one for diluent, and allows the sample to be aspirated, diluted, and dispensed with one fire of the trigger. The diluent chamber can be removed easily to change reagents quickly, according to Labsystems, and is autoclavable for sterile applications. The £249 (UK) Diluter has a volume range of 0–100 μl for samples and 90–1,000 μl for reagents, and can prepare dilution ratios of between 1:1 to 1:200.



Shoots dispensing worries.

To satisfy laboratory needs associated with the increased use of radioimmunoassays (RIA), V.A. Howe & Co. Ltd. distributes a line of **Hamilton diluter/dispensers** (Reader Service No. 115). The £3,370 (UK) Microlab MT is a microprocessor-controlled four-channel dispenser designed for microtitre plate work used in ELISA. V.A. Howe also sells the smaller Microlab M for routine work, and the Microlab 1000 for high dilution ratio (1:5,000) applications.

Erratum: Zeiss microscopes

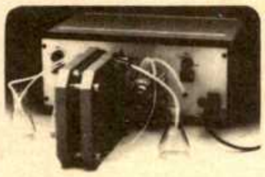
A review published in the 3 July issue of *Nature* (**322**, 93) incorrectly stated that "budget priced" Zeiss microscopes were available from Gallenkamp. While Gallenkamp does supply a wide range of Zeiss microscopes, the "budget" microscopes mentioned are not manufactured by Zeiss, and are badged Gallenkamp.

ADVERTISEMENTS

LIPOSOMES

NEW

REPRODUCIBLE – HOMOGENEOUS – EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM
POB 126, D-8000 Munich 65, FRG

Reader Service No.35

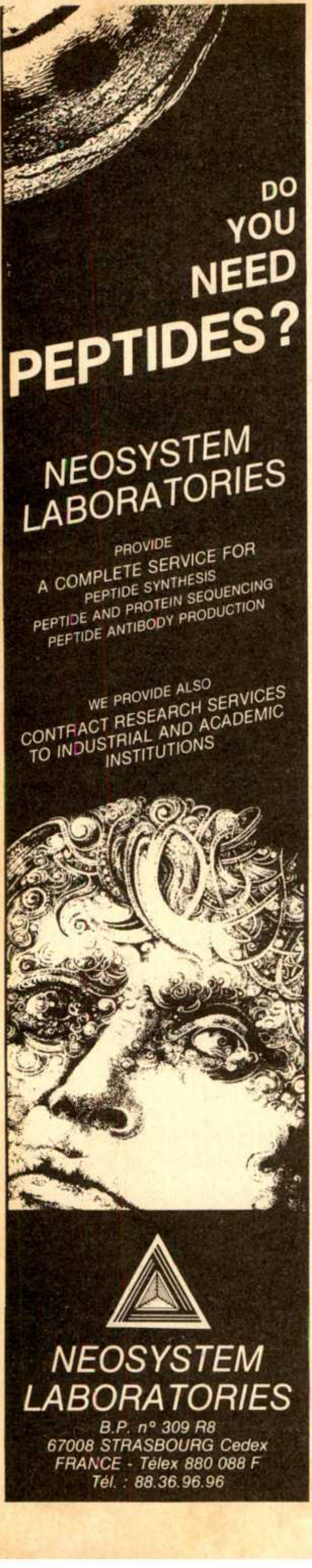
Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

Low Cost

Eliminates evaporation loss and contamination. Easy to use – flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. **billups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.**

Reader Service No.7



ADVERTISEMENT

Roland GRUNBERG - NANCY - FRANCE

DO YOU NEED PEPTIDES?

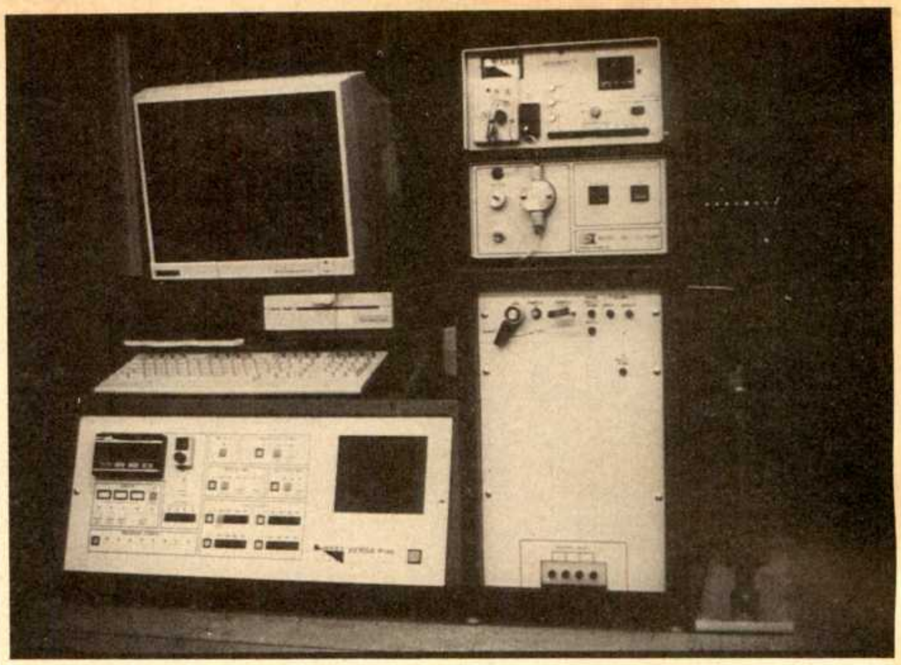
NEOSYSTEM LABORATORIES

PROVIDE
A COMPLETE SERVICE FOR
PEPTIDE SYNTHESIS
PEPTIDE AND PROTEIN SEQUENCING
PEPTIDE ANTIBODY PRODUCTION

WE PROVIDE ALSO
CONTRACT RESEARCH SERVICES
TO INDUSTRIAL AND ACADEMIC
INSTITUTIONS

NEOSYSTEM
LABORATORIES
B.P. n° 309 R8
67008 STRASBOURG Cedex
FRANCE - Télex 880 088 F
Tél. : 88.36.96.96

Reader Service No.4



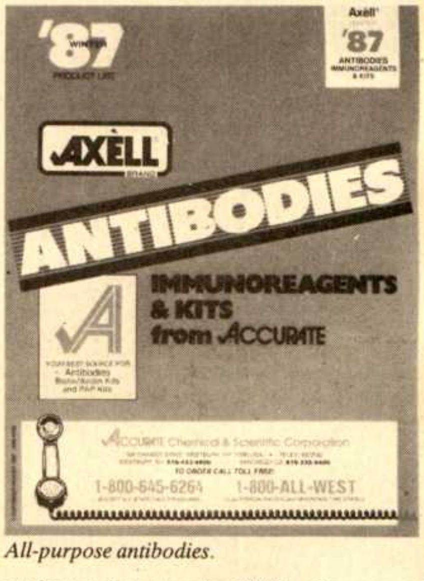
The VERSAPrep-GT grows with you.

The VERSAPrep series of **modular preparative HPLC instruments** offered by Varex Corp. allows researchers to expand according to their specific needs (Reader Service No. 116). The basic VERSAPrep-GT costs $25,465 (US) and may be operated at flow rates of up to $180\ \mathrm{ml\ min^{-1}}$ at a maximum pressure of 3,000 p.s.i. according to Varex. Samples may be introduced with a valve-loop injector, the main eluent pump, or an accessory analytical-scale pump. Varex's HPLC offers a variety of accessory modules. A Control/Automation Module fully automates the basic VERSAPrep, while the Fraction Collection Module adds four collection ports to the VERSAPrep-GT's capabilities. The Programmable Events Module permits complete computer control, and the Column-packing Module allows the operator to slurry pack columns up to one inch in diameter. Varex says the precolumn module minimizes elution time by trapping extraneous, late-eluting components, and can be back-flusheda with solvent. The introduction of the Multi-Column module expands the VERSAPrep by making it possible for two columns to be selected, either individually or sequentially, as required for a given run.

Catalogues etc.

Pharmacia is giving away a **free wall poster** describing chromatographic techniques for the purification of monoclonal antibodies (Reader Service No. 117). The 70 × 100 cm poster is designed as a useful as well as decorative addition to a laboratory, and features charge-specific, FC-region specific and immunospecific methods for chromatography and FPLC. Pharmacia's new catalogue of separation products for immunology includes information on products for cell affinity chromatography, DNA synthesis, cell culture, cell separation by centrifugation, analytical and preparative electrophoresis systems, monoclonal antibody purification and cell studies.

Accurate Chemical has recently published a **catalogue** of its Axell Antibodies (Reader Service No. 118). The 84-page catalogue lists nearly 3000 primary and secondary antibodies, including ready-to-use diluted primary antibodies and negative controls. Accurate Chemical also provides a selection of immunochemistry kits and reagents.


'87 WINTER
PRODUCT LIST
Axell '87
AXELL
ANTIBODIES
IMMUNOREAGENTS & KITS from ACCURATE
ACCURATE Chemical & Scientific Corporation
1-800-645-6264
1-800-ALL-WEST

All-purpose antibodies.

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

nature

Arms, hostages and petroleum

The inquiry into the dealings between the United States and Iran has so far concentrated on the supply of arms. What promises have been made about the price of oil?

PRESIDENT Ronald Reagan's discomfiture at the revelation in the past two weeks of his secret and possibly illicit dealings with the government of Iran is understandable. The proposition that the United States has a natural interest, but also a kind of public responsibility, to cultivate good relations with seemingly hostile governments makes sense so far as it goes, and might be regarded in the Congress as justification for an attempt at rapprochement with Vietnam, North Korea or (come to that) the Soviet Union. But Iran? With the memory of the capture of the staff of the Tehran embassy's staff still so fresh? That alone would expose Mr Reagan to the complaint that deals with hostage-takers can serve only to put further potential hostages at risk. So much is so clear that it is inevitable that people should believe what the president denies, that the "small shipment" of arms to Iran was a consideration for Iranian influence with the captors of several US hostages still held in the anarchic Lebanon. No doubt the tail-enders of the 99th Congress will get nearer to knowing what really happened.

Meanwhile, there is a larger question to be explored than that of whether some part of the US administration was prepared to make a deal most other people would have considered unwise. When the news filtered from Iran at the beginning of this month that Mr Robert McFarlane, director of the US National Security Council until a year ago, had taken to Tehran not merely a small shipment of arms, but a bible and a cake, the word was that the United States was also willing to use its influence to bring about a general increase in the price of oil, from something like $15 a barrel now to something like $18. While attention centres on the arms and the bible (it is harder to imagine what symbolic purpose the cake may have served), the price of oil tends to be overlooked. yet on the assumption that dealing over hostages will now become unfashionable, this is a more lasting and more serious issue.

Circumstantial evidence suggests that something odd may have been going on. As a member of OPEC (the Organization of Petroleum Exporting Countries), Iran has recently vociferously expressed its discontent at the fall by more than a half, even in nominal terms, of the price of oil on the international market. Its own income from petroleum has been cut not merely on this account, but by interruptions of the movements of oil tankers from its terminals at the head of the Persian Gulf. But for the past several years, as for most of the time since OPEC's beginning in 1969, the market in oil has been determined by the policies of Saudi Arabia, whose reserves of cheap oil are now even more dominant than before 30 years of vigorous exploitation in the Middle East. Apparently exasperated by OPEC's failure to agree on production quotas tight enough to sustain a higher price, Saudi Arabia has for the past two years gone for market share instead, increasing its own production and letting the price find its own level to the general benefit of the industrialized West and the discomfort of relatively high-cost producers of oil, Iran for one and Britain (not a member of OPEC) for another.

That policy has now been turned on its head. Sheikh Ahmad Zaki Yamani was dismissed his post as oil minister of Saudi Arabia a month ago, since when his successor has arranged a meeting of OPEC advertised as one at which the price of oil should be increased. There are several explanations, of which one is that the United States has successfully exerted its influence with the Saudis, in the manner suggested by the shadowy (but at least partly accurate) sources in Tehran. It goes without saying that this sequence of events affecting the price of oil may be as plausibly explained by Saudi exasperation that the United States had begun dealing with Iran again; Saudi Arabia may even have resolved that, if anyone deals with Iran, it should do so itself. There are also many other explanations in which President Reagan's little shipment of arms plays no part.

For the time being, the US administration's interest is that it should if possible be cleared of the suspicion that it has been prepared to monkey with the price of oil on the international markets. In its way, this charge is as serious as is the charge that the United States has been prepared to trade with those who support or encourage kidnappers, thereby helping to institutionalize political hostage-taking. Not that Iran would be the only beneficiary of even a modest increase of the price of oil. The British chancellor of the exchequer would benefit, as would his opposite numbers in Nigeria and Mexico (not to mention those who worry about Mexico's debts). So too would US oil producers in Texas and California. But the gret majority of those who buy OPEC oil, from Japan to Western Europe, would be hurt by an increased price of oil, while the flickers of hope that economic growth can yet be rekindled would become more fitful. These are the players who will want to know what, if anything, President Reagan and his emissaries have been telling the oil producers of the Middle East. □

Insiders outlawed

Cleansing the world's financial markets of fraud may be more difficult than supposed.

THE wave of righteous indignation about insider trading on Wall Street and, on a lesser scale, in London, is understandable but suspect. Not that those who use privileged information to play the stock markets to their own advantage are distinguishable from common thieves. Mr Ivan F. Boesky, the New York financier required by the US Securities and Exchange Commission two weeks ago to hand over $50 million in ill-gotten gains and to pay the same amount as a penalty, is the most spectacular case so far. But the case of Mr Geoffrey Collier, the director of Morgan Grenfell in London who was forced to resign his job after seeking to make £15,000 on a single insider trade, is more illuminating if only because it is easier to understand.

Collier's fault was to have sought to buy shares in a British engineering company in advance of a takeover bid of which he had private knowledge, gathered in the course of his work. The deal was executed on his behalf by an overseas company and through a New York stockbroker, which accepted the order to supply the stock at an agreed price even though it did not have enough on its books at the time. The stockbroker's calculation was that it could have bought the stock in the markets later in the day, but in the event it laid off two-thirds of the risk that the price might rise onto another New York firm, a subsidiary of Chase Manhattan. If the deal had not been cancelled, one firm would have lost £5,000 and the other £10,000.

So what? will be one cynical question. Stockbrokers are forever making huge sums of money by "dealing their books", so why weep when they risk a loss? The simple answer is that it is not their money that they risk, but ours, that invested in the international stock markets by pensions funds, savings institutions and by banks that take deposits from the general public. If Collier had pocketed his £15,000, the ultimate losers would have been ordinary people. Boesky's $50 million ultimately derives from the same source. But $50 million is only a tiny fraction of the funds belonging to the public invested in the international financial markets. Is the rest safe?

What worries the Securities and Exchange Commission now is that, in the United States, the huge amount of paper debt obligations called "junk bonds", perhaps as much as $100,000 million, which has sustained the recent wave of spectacular takeovers, may be suspect. These securities (if that is the right word) entitle the owner to high interest, but are not secured against assets and not traded on stock exchanges. Holders can get back their capital by appealing to the investment banks that issued the junk bonds, but the fear is that this procedure may have been infected by Boesky get-rich-easy principles; that could spell disaster, given the extent to which savings and loan associations in the United States have traded their clients' cash for junk bonds (which look better on a balance sheet).

So there is a strong case for making outlaws of the insider-traders. How is that to be done? This is the point at which righteous indignation comes a cropper. The case of Collier is straightforward; he acquired his inside knowledge because the firm of which he was a member had agreed to mount a bid for the British engineering company (on behalf of Mr Robert Maxwell, the proprietor of Pergamon Press). To seek to profit personally from this knowledge was a violation of his firm's internal regulations, of the London Stock Exchange's rules and of British government law (hurriedly brought into effect in time to put Collier under pressure). Boesky's is a more complicated case; so far, he has been accused only of using inside information provided by a true insider, Mr Denis Levine, arrested by the Securities and Exchange Commission earlier this year and who has since traded immunity from further prosecution for the repayment of $1.5 million and continued cooperation with the commission's inquiries; who, in law, is the real felon, Boesky or Levine?

The truth is even more complicated. Without accepting the arguments of the Chicago economists who say that insider trading is virtuous because it brings reality more speedily into effect and thus make the markets more efficient, it is fair to ask whether stockbrokers who commission studies of industrial sectors, ostensibly for the benefit of clients, should then be allowed to speculate in the stocks concerned. And should the 'analysts' who carry out the research be allowed to speculate on their own accounts, knowing as they must that at least some of their company's clients will believe what they say? When is it improper to act on (and profit from) rumour? Or even to take advantage of special knowledge of the health (or sickness) of publicly-quoted companies gathered in quite different connections, as a journalist, for example?

The righteously indignant will quickly say that there should be no limits to the most simplistic writ. (Simplistic Senator William Proxmire is alarmed at the way in which corporations have been saddled with huge amounts of debt after successfully defending against a takeover bid, but that is another issue for which directors are responsible.) Others will recognize that there is an urgent need for a legal definition of a boundary between right and wrong that will make sense, and will endure. Governments, poor things, will know that the problem is complicated by the international character that the financial markets have now acquired. □

Teachers in travail

Is the proper education of the young too important to be left to teachers?

Like motherhood in the old days, the education of the young has become an important social objective, and for good reason; economic survival depends on social or technical innovation, which requires innovative young people. So it is not surprising that governments in the industrialized West should be seeking to improve high-school education. In the United States, the new Democratic Congress has lost little time in saying that it wants the federal government to do more for state-supported school systems. (What will happen to the Gramm–Rudman budget deficit reduction act if the defence budget is not cut?) In Britain, the government is in the last bruising stages of an attempt to trade money (which teachers need) for a promise (which teachers' labour unions are reluctant to concede) that more able teachers will be better rewarded. Everywhere, the simple arithmetic is overlooked.

Here is the arithmetic. Most Western countries send their children to school for 12 or 13 years, in classes that average about 20 students per teacher. In round numbers, the schooling of each child requires roughly two-thirds of a teacher-year. But teachers' professional lives last for about 35, and are statistically much shorter, which means that for a demographically unchanging population, at least 2 per cent of the population of working age should be schoolteachers. Naturally, governments and the societies they represent require that teachers should have had the benefit of higher education, in which the participation rate in Britain (admittedly one of the worst) is merely 15 per cent. If all those leaving higher education followed professional careers for 35 years (which is far from being true), one seventh of the population would have to be schoolteachers. Allowing for the manpower needs of higher education, one should expect that roughly 20 per cent of college-leavers would be occupied in education. Is it reasonable to expect that those who teach the young will also be among the most talented?

There are three distinct answers to this question. In the United States, where the federal government has, for the past six years, stuck largely to exhortation, but where the quality of high-school education has become a general anxiety, some state governments are making valiant efforts to improve the performance of their schools, while institutions that train teachers are trying to enhance the reputations of those who teach in schools by requiring more from entrants to their courses. These private initiatives have not so far been well-rewarded because people will not pay particularly well-qualified teachers extra, while teachers themselves pretend that discrimination is anathema. In Britain, on the other hand, the government is willing to battle for its conception of a core curriculum and is also pleading for pay-discrimination; it should be known this week or next whether its offer of an extra 17 per cent over 18 months will be accepted on the teachers' terms or will have to be enforced. Only in Japan is there is a workable if unstable solution of the problem: put more than a third of the population through higher education, but then arrange that married women do not work, but stay home and teach their children.

Elsewhere, governments will have to learn to live with solutions far from their ideals. In North America and Western Europe, the chance of persuading the best and brightest of those leaving higher education to teach in schools is small, especially in those fields from which technology springs. Better to be reconciled to that than to keep tilting at a windmill. So the crying need is for a general recognition that schooling is not an end in itself but one of several means to functional adulthood, that the core curriculum should be stripped down so that it is no longer a test of the preternatural ability of gifted students but within the competence of all students and most teachers and that, whatever governments do, the results will not be perfect. □

Europe's framework research

Commission ready to pull out communal rug?

Brussels

THE European Commission could go so far as to drop the whole $8,000 million, 'framework five-year programme for European research and development' next month, if governments fail to agree to support the programme at a sufficiently high level. So observers were saying in Brussels last week, as the Commission began to field its strongest political musclemen, Commission President Jacques Delors and his team, in favour of what the Commission considers a sensible budget for the research programme.

Last week, the European Parliament's Energy, Research and Technology Committee voted to advise the Commission to withdraw the framework programme if the full budget requested — a reduction compared with the $10,000 million earlier requested and a "minimum" according to the committee — was not accepted.

Certainly research, which at the full amount of the framework programme would account for little more than a twentieth of Brussels spending, is now at the centre of Brussels politics, where it is feeling the full blasts of European summitry. The Commission wishes through research to confirm a new role for itself in the industrial regeneration of Europe. France, West Germany and the United Kingdom have been resisting both for financial reasons and, undoubtedly, because of the interference in national science and technology programmes that increased Commission activity would imply.

In fact, governments have accepted the principle of a major European programme of research and development coordinated by the Commission in the 'Single Act', the major treaty of union agreed by heads of state last year in Luxembourg and due to be voted in place by all members of the European Communities by the end of this year. The Single Act includes a place for the framework programme for research, and Delors sees the voting and proper financing of that programme to be a clear test of the 'New Europe'. He told British industrialists recently that European Communities research and development policy was "so relevant and so essential" because in the long term it bore directly on the creation of a single European market for European goods.

The framework amounts to less than 2 per cent of overall research and development in Community countries and "we are not, repeat not, out to enter into some kind of competition with the member states", Delors said. "We hear references to cost-effectiveness. But the microchip industry people tell us that their return on ESPRIT programmes is 400 per cent despite the cost of cooperation. We are said to be bureaucratic. But do people know that it takes two European officials to manage £1 million of research and development funds, and eight officials in the most efficient member states? Industrialists, and researchers, and academics are turning to us, to Europe, and saying they are ready to work together... I do not see why the most solemn undertakings by the heads of state and government [in the Single Act] are not being put into effect faster."

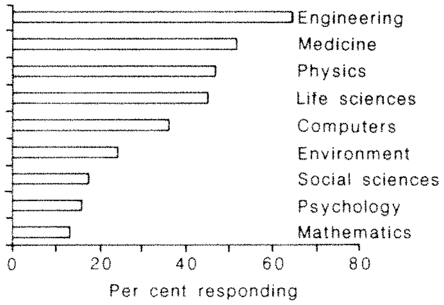
Robert Walgate

US universities appear in good shape

Washington

CONSTRUCTION in progress and planned over the next five years will increase research space at US research universities by between 14 and 26 per cent, according to a survey* by the National Science Foundation, at a cost of $7,500 million.

The survey covered all doctorate-granting institutions, and showed that the top 50 in terms of research and development spending (out of a total of 165) accounted for more than 50 per cent of work in progress and 60 per cent of planned work. Engineering, medical sciences and physical sciences departments were more likely to have construction in progress than others (see figure).



The survey produced one surprise for the doomsayers on university research facilities: over half of the same top 50 universities reported the condition of their research facilities as "good" or "excellent".

But of those ranked outside the top 50, over half replied "fair" or "poor". ☐

**Science and Engineering Research Facilities at Doctorate-granting Institutions*. National Science Foundation, 1986.

UK research appointments

Council breaks new ground

THE terms on which the Natural Environment Research Council (NERC) has made three senior appointments to its staff have occasioned some surprise among British academics.

Interest centres particularly on the arrangements under which the three directors of science appointed in the past twelve months have been promised that, if their five-year terms are not renewed, they will be kept on the payroll at a professional level.

The three appointments were foreseen in NERC's corporate plan, published nearly two years ago. The objective was that NERC's spending in three areas (Earth science, marine science and terrestrial and freshwater ecology) spanning its whole range of interests should be overseen by scientists of individual distinction. When the new arrangements were announced, some directors of NERC institutes expressed misgivings on the ground that their independence would be compromised and their access to the NERC council impeded.

The first appointment, of Dr P.B.H. Tinker (terrestrial and freshwater ecology) was made last December. Dr J.D. Woods (marine science), previously at the Kiel institute of oceanography, and Dr J.C Briden, previously professor of geology at the University of Leeds, were appointed in the summer of this year.

As part of the arrangement with the holders of the three posts, it has been agreed that they should remain active in research, for which reason it is intended that they should maintain modest research programmes at suitable academic institutions (All three have plumped for the University of Oxford, within driving distance of NERC's Swindon headquarters.) Dr J.C. Bowman, secretary of NERC, says on the telephone that the council of NERC has agreed that sums "of the order of 100K [£100,000]", taken from the budgets of the three directorates, should be used to support this research.

On NERC's promise of continued employment, Bowman says that the arrangements do not differ from those governing the appointment of scientists elsewhere in the council's network, who are normally offered employment until retirement age. Moreover, he says that the arrangements are not inconsistent with the British Treasury's rules on public appointments, but merely a "new wrinkle" on them.

Tinker will be 60, the normal retirement age from the public service, when his five-year appointment ends, but both Woods and Briden (now 47) will have some years to go.

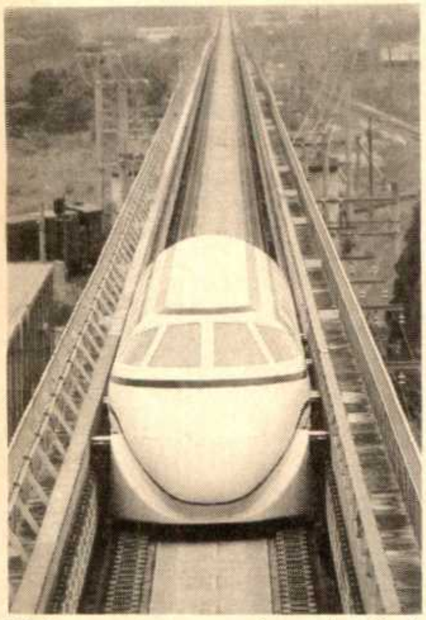
John Maddox

Magnetic levitation

Japan's railways look ahead

Tokyo

HIGH-SPEED magnetically levitated (Maglev) trains are in the news again. When Japan, which has been running a modest Maglev research programme, invented a



517 km per hour experimental train in 1979, few thought it could ever emerge from the test track. But the climate has suddenly begun to change and politicians are openly talking of a magnetically levitated bullet train that could transport passengers from Tokyo to Osaka in little more than an hour.

A prototype train that can carry up to 80 passengers is now under construction. Wind tunnel tests on a 7-m scale model of the prototype were successfully carried out at Nissan Motor Company's Technical Centre two weeks ago, and the prototype itself should be ready for test runs on Japanese National Railways (JNR)'s 7-km test track in Mihazaki, Kyushu, next spring.

The break-up and privatization of JNR has been the catalyst. Years of profligate government spending and intransigent unions have left JNR with debts of $230,000 million, more than the combined national debts of Mexico and Brazil. Interest payments alone have been running at more than $20 million a day.

Eight bills have raised through the Lower House of the Diet tailored to break up JNR into six regional passenger railways, a freight railway and a holding organization for the profitable bullet train lines (Shinkansen) which will be leased back to the passenger railways. More than 3,000 hectares of JNR land, much of it in prime areas, will be sold off to recoup some of the debt and 61,000 JNR employees will have to find employment elsewhere. Yet apart from a few radical attacks on the railways communication system which brought temporary chaos to commuters (see *Nature* **318**, 405; 1985), opposition to the privatization plan has failed to materialize and many union members have capitulated for fear of losing their jobs.

After the break-up in April 1987, JNR's research laboratories will be reborn as a research institute backed by a foundation. More money is to go into Maglev. Maps for new Shinkansen lines, pencilled in during Japan's high-growth era, are being taken out of the back drawers of filing cabinets for reappraisal. It is hoped that one of these lines could carry a Maglev train.

The Maglev train floats on the magnetic field of superconducting niobium–titanium magnets mounted below the car floor. As the magnets pass over aluminium coils set in the U-shaped trackway, electric currents induced in the coils set up a repulsive magnetic field off the same polarity that levitates the vehicle about 10 cm off the ground. Guidance and propulsion are provided by electromagnets set in the walls of the trackway.

The train offers a smooth, silent and bump-free ride at more than 400 km per hour, and in commercial service will be controlled by ground stations at 50-km intervals along the track. One slight problem, however, remains. In an earlier test, the maximum magnetic flux on the floor of the car was 200 gauss. Passengers complained that their watches stopped. In the prototype the disposition of the train's superconducting magnets will be adjusted to try to overcome this effect, according to Yoshihiro Kyotani, manager of the Maglev project.

Kyotani already has the most likely route for the first commercial Maglev line dotted on a map in his office, the Chuo (Central) Shinkansen, which would run from Tokyo to Osaka via Nagoya.

Although nobody is yet discussing construction costs, Kyotani points out that the trackway for the new vehicle would cost less than that for a conventional railway. But he admits that the cost of the Maglev vehicle itself will probably be higher than that of an ordinary train.

David Swinbanks

SERC scoops pool

EARLIER this week, it seemed certain that the Science and Engineering Research Council (SERC) would be given the lion's share of the extra £25 million for research provided three weeks ago by the British government. This appears to have been the

CERN
PLEASE HELP SERC. THANK YOU

decision of a meeting last week of the independent members of the Advisory Board for the Research Councils (ABRC), the government's cake-cutting committee, and is unlikely to be overturned by a meeting of the full committee arranged for Wednesday this week.

SERC's success is a pyrrhic victory. The extra funds channelled in its direction will be used exclusively to pay the increased cost (in sterling) of subscriptions to overseas research organizations, among which the European high-energy physics research centre at Geneva (CERN) is the largest. The cost of the CERN subscription, now roughly £45 million, has risen because of the devaluation of sterling relative to the Swiss franc. □

Soviet reservations on IAEA conventions

THE Supreme Soviet has approved the two new conventions of the International Atomic Energy Agency (IAEA) formulated during the summer in the light of the Chernobyl catastrophe. The signing is hardly remarkable; the effect of Chernobyl was to crystallize ideas that had begun to take shape among IAEA members after the accident at Three Mile Island. But Soviet media and government spokesmen have consistently claimed the new conventions as a specifically Soviet initiative.

But the Soviet Union has ratified the treaties only in part. A special reservation has been added to the effect that the USSR will not consider itself bound by the provisions of Article 11.2 of the Convention on Prompt Notification of a Nuclear Accident nor by Article 13 of the Convention on Assistance in the Event of a Nuclear Accident or Radiation Emergency. These provisions refer to the possibility of submitting any dispute arising under the terms of the conventions to arbitration or to an international court at the request of any side. The Soviet Union, says the reservation, declares that the consent of all sides is necessary before any dispute is referred to arbitration or to an international court.

Vera Rich

US polar research

NSF break with Coastguard?

Washington

PLANS for polar research have been upset by the US Coastguard's recent downgrading of one of its ships, *Glacier*, which is no longer classified as an "icebreaker", but described as "ice-strengthened" instead. As a result, the National Science Foundation (NSF) may break with the US Coastguard, and find polar research vessels elsewhere.

Glacier was considered the most suitable of the Coastguard's icebreakers for research, and the discovery that its hull has been seriously weakened by corrosion has re-opened the debate about whether Coastguard ships are still suitable as research platforms for NSF. Two new Coastguard icebreakers should be completed in 1992 or 1993, but NSF will have to span the gap from now until then.

The nub of the problem is that Coastguard ships are primarily military vessels not well equipped for research. Two years ago, NSF leased one small research icebreaker, *Polar Duke*, and is also now taking research space on the West German icebreaker *Polarstern*, which has superior research facilities. But restoring *Glacier* to icebreaker status would cost $20 million and NSF considers the cost would be too great.

Instead, officials want the use of a dedicated research icebreaker from another source, either by collaboration with another country or by leasing. Without NSF support, the Coastguard cannot justify large expenditures on *Glacier*, although it says no final decision has been taken. Loss of the NSF account would mean about $2.3 million a year.

Scientific interest in Antarctica seems to be increasing year by year. Investigation of the Antarctic ozone hole seems now to guarantee that interest will not wane. NSF, as well as using *Glacier*, has also frequently used two other Coastguard icebreakers, *Polar Star* and *Polar Sea*. But these are unpopular with researchers and in need of refits.

Glacier is working this season in the Antarctic ice edge zone, pushing but not breaking ice. Research has not been unduly affected this year, but it could be if NSF were tied to the ship. The two new vessels due next decade will provide better accommodation and facilities than any existing Coastguard ship. Meanwhile, NSF is looking at its options. **Tim Beardsley**

First leptonic collisions at Japan's KEK

Tokyo

TRISTAN, the giant accelerator at Japan's High Energy Physics Laboratory (KEK), burst into action last week when electrons and positrons were smashed together at a record-breaking centre-of-mass energy of 50 GeV in the accelerator's 3-km main storage ring. And with foreign competition breathing down its neck, the TRISTAN team is now racing to get its experimental programme into full swing.

After successful tests of the electron and positron beams at 25.5 GeV earlier this month (see *Nature* 324, 7; 1986), KEK scientists set the counter-rotating beams in the main ring on collision course and the first hadronic-like event was picked up by the partially instrumented VENUS detector shortly before midnight on 18 November, followed a few hours later by a collision producing large-angle Bhabha scattering.

The record set by TRISTAN exceeds the maximum of 47 GeV achieved by the PETRA ring in Hamburg. But the maximum luminosity obtained, 2×10^{29} cm^{-2} s^{-1}, is barely a hundredth of TRISTAN's design capability, and the electron- and positron-beam currents, at present operating at around 1 mA, will have to be raised before a serious search for new elementary particles can begin.

Aerial view of the TRISTAN complex.

AMY, a detector built in collaboration with teams from the United States, Korea, China and Japan, has just been rolled into position in one of the four experimental halls on the main ring. And the TOPAZ detector, now undergoing tests with cosmic rays, should be in place next spring ready for the full experimental programme to begin in early May 1987.

TRISTAN thus has a head start on its most immediate rival, the Stanford Linear Collider, whose even higher energy electron–positron collisions are not expected for at least a few months. Even then, the linear collider will be limited to a single energy, producing showers of intermediate vector bosons (the Z^0), TRISTAN will have about two years as the world's most powerful flexible electron–positron ring collider until the LEP ring at CERN comes on line in 1988. David Swinbanks

---

Optical interferometry

Large grant for new instrument

Sydney

THE University of Sydney has won a grant of A$500,000 to begin construction of a 640-m baseline high-resolution stellar interferometer. This development is a milestone for stellar astronomy (and for the Chatterton department of astronomy at Sydney) as well as for the Australian Research Grants Scheme, whose budget is so small that projects on this scale have so far been beyond its scope.

The new instrument will be a larger version of a prototype instrument built by Sydney over the past 15 years. Only about 30 stars, mostly bright giants, have so far had their apparent diameters measured to an accuracy within 5 per cent by this technique. The new interferometer telescope will be able to measure the apparent diameter of 50,000 stars (brighter than 8th magnitude) with an accuracy of 2 per cent or better.

Measurement of the apparent size of stars whose distance is known should allow the establishment of fundamental scales of stellar diameters and luminosities to be established. Observations of close binary stars can give all the parameters of the system, including their orbits, temperatures, luminosities, diameters and distances. Other problems the interferometer could help solve include mass loss, by observing how a star's envelope appears to grow at longer wavelengths, and stellar rotation, by observing the flattening of the poles.

All the essential technology for the interferometer telescope has been proved in a prototype developed by the Sydney University team, led by Dr John Davis, over the past 10 years. The prototype has an 11.4-m baseline with a steerable mirror at either end. The mirrors capture the light from a star which is then reflected to a central optical bench where the beams are combined and interference fringes observed. The prototype will form the heart of the new instrument.

The principle on which the new interferometer will operate is amplitude interference, its design being a development of Michelson's classic interferometer. Amplitude interferometry has not previously been a workable option because of the blurring caused by atmospheric 'seeing' effects. Until this year the Australian Research Grants Scheme has been restricted to grants of less than A$150,000. The $150,000 grant for the interferometer marks the beginning of the government's policy of more funds for the scheme concentrated into fewer but more highly rewarding grants (see *Nature* **323**, 98; 1986).

Charles Morgan

Environmental databases

NOAA slipping further behind

Washington

THE attempts by the US National Oceanic and Atmospheric Administration (NOAA) to bring up to date its overstretched environmental data archiving and distribution system have run into trouble. The agency's proposed new system, NOAANET, is intended to supply retrospective environmental data from satellite and other sensors for researchers during the 1990s and into next century. But political obstacles have prevented NOAANET from finding its way into the agency's budget for fiscal year 1988, which starts next October, and there are fears that the administration's outmoded technology could slip even further behind.

NOAANET was cut out of the agency's 1987 budget by the Office of Management and Budget. This year, it did not even get that far: NOAA has been told by its parent agency, the Department of Commerce, to see whether the system could not be run by the private sector. Nonplussed NOAA officials say prospects for making a profit out of environmental data archives are not immediately clear.

Virtually all researchers agree that NOAA has fallen way behind in the push to computerize access to environmental data, even though it has a legal mandate to provide such an archive. Requests for data can take months to be processed, and there is no on-line catalogue, let alone data access. The problem will worsen as new advanced satellites will generate data in ever-increasing quantities. The National Aeronautics and Space Administration (NASA), in contrast, although supposedly a research agency, has already established data centres with much more advanced capabilities and is starting to link them up. The new 'NASAnet' will enable researchers to examine catalogues of data from different sensors across various disciplines and make selections from their computer terminals.

The difficulties result, some say, from NOAA's disadvantaged position within the Department of Commerce. The department's bureaucrats are less likely to be sympathetic to projects benefiting research than to those of NASA, for example. NASA's initiatives to make use of computers for data handling and distribution have been generously funded: the agency recently signed a $240 million contract with Boeing Computer Services for project support for a new network.

There are nevertheless those who are critical of NOAA's preliminary plans for NOAANET. Frank Eden of the Joint Oceanographic Institution questions whether there have been adequate consultations with NASA and the National Science Foundation, for example, which is also constructing a large computer network based around its five supercomputer centres.

The lack of a formal scientific advisory body to advise NOAA on how to implement the project is also a source of concern. But Gregory W. Hunolt, NOAANET project manager, says NOAA plans to bring in external advice shortly; meetings with NASA have already started and a definition document has been circulated in the academic community. High-level contacts between the agencies had been delayed by a NOAA proposal to put its orbiting meteorological instruments on a private satellite, Omnistar, but that idea has now been dropped. But without a commitment by the Department of Commerce to seek the several tens of millions of dollars that NOAANET would cost, progress seems unlikely.

Tim Beardsley

AIDS

WHO plans megaprogramme

Washington

THE World Health Organization (WHO) last week said that prevention and control of AIDS (acquired immune deficiency syndrome) has become one of its top priorities and that efforts to combat the disease need to be as urgent as those against smallpox in the 1950s. WHO estimates there will be between 500,000 and 3 million AIDS cases by 1991, with up to 100 million people infected with the virus.

Dr Jonathan Mann, head of WHO's AIDS programme, says those figures could be serious underestimates if AIDS spreads quickly in Asia and in South America. WHO estimates the worldwide cost of fighting the disease at $12,500 million by the eary 1990s.

The WHO AIDS programme officially started in May of this year, and has already raised $4.5 million in special contributions from member governments. The agency is seeking to raise $200 million for next year. By next decade it wants $1,500 million a year, 50 per cent more than its entire budget at present.

WHO will concentrate on the exchange of information among researchers working on prevention and therapy, as well on direct assistance to countries needing to set up their own control programmes. The research component will include coordination of vaccine field trials and the establishment of an international virus bank.

Mann estimates that about 100 countries will need technical assistance in setting up their own prevention and control programmes. These should be based on research on how the disease is transmitted and public education, and should include "hard-headed" social science research into the availability and usage of condoms and hypodermic needles, for example. WHO will also foster operational research into the development of cheap antibody testing for blood banks worldwide.

Tim Beardsley

- The British government last week launched a £20 million campaign to educate the public on the danger of AIDS. "Don't die of ignorance" is one of the slogans that are to confront the British public in a wave of publicity including television, radio and maildrop advertising. The campaign began this week with national newspaper advertisements.

At the end of the year, a month-long television campaign using a 40-second commercial will pave the way for the maildrop. In the new year the 23 million homes in the United Kingdom each receive a letter containing an explicit warning on the danger of AIDS.

A minute-long radio commercial will also be broadcast at the end of the year, narrated by a top disk jockey to attract the attention of the young. **Bill Johnstone**

India screens foreign students for AIDS

New Delhi

MEDICAL screening for AIDS has been made compulsory for all foreign students seeking admission to Indian educational institutions. Under the new guidelines from the health ministry, foreign students provisionally admitted may be liable for deportation if tests confirm the presence of antibodies to human immune deficiency virus (HIV).

The government directive was issued last month after the detection of HIV antibodies in an African student in Madras. He is the only male seropositive case among 22 cases so far identified in the city.

India has an estimated 30,000 foreign students, mostly from Africa and eastern Asia. AIDS screening has been proposed for fresh admissions, but is likely to be made mandatory for foreign students already here. The Indian Council of Medical Research has set up 19 centres to screen sera using imported ELISA (enzyme-linked immunosorbent assay) kits.

There have so far been three deaths from AIDS in India — two of the patients had received blood transfusions in the United States and one was an Indian settled in Malawi. Some 30 antibody-positive cases, most of them prostitutes, are under observation.

K. S. Jayaraman

Japanese volcano

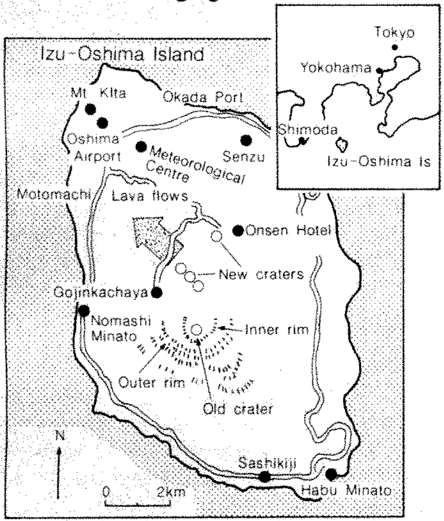
Islanders flee surprise eruption

Tokyo

JAPAN's volcanologists were taken by surprise when Mount Mihara on Izu-Oshima island south of Tokyo exploded in a series of violent eruptions last Friday, forcing the overnight evacuation of the island's 10,000 inhabitants as a river of lava threatened to engulf the main town.

Izu-Oshima is the first in a chain of volcanic islands that extends southward for more than 1,000 km into the Pacific along the edge of the Philippine plate. A submarine volcano at the southernmost end of the volcanic arc temporarily broke surface during an eruption earlier this year (*Nature* **319**, 441; 1986), and Miyakejima, Izu Oshima's immediate neighbour, erupted in 1982. But Izu-Oshima itself has been silent for the past 12 years, and an eruption on the present scale has not occurred in the past two centuries.

The first signs of activity came in July when the island was rocked by volcanic tremors. But surveys by the Meteorological Agency and Tokyo University's Earthquake Research Institute showed no evidence of the 'bulging' of the Earth's crust that typically precedes major eruptions as magma rises to the surface. And at the end of October the Volcanic Eruptions Prediction Liaison Council declared that there were no signs that a large-scale eruption was imminent.



Even after eruptions began on 15 November, the council confidently predicted that lava flows would be confined to the caldera by the outer rim of the volcano, and Oshima enjoyed a temporary boom as hundreds of tourists flocked to the island each day to enjoy the spectacular eruption and lava flows.

After 10 million cubic metres of lava had spilled over the inner rim of the volcano and the rivulets of lava in the caldera had frozen in their path, the Meteorological Agency announced that activity appeared to be decreasing and the island's residents were in no immediate danger. Hours later the evacuation order came.

What nobody had foreseen was the opening of a new line of vents that cut through the outer rim, something that had not occurred in more than 500 years. As the new vents burst into life, lighting up the night sky in a curtain of fire, a river of lava set forests ablaze as it surged towards the main town of Motomachi.

By dawn the island was deserted, the 10,000 residents and 2,000 tourists having been evacuated to the mainland. All that remained were abandoned domestic animals, a farmer who stubbornly refused to leave his 5,000 chickens, and some baffled volcano experts.

On Monday, the lava front lay within a few hundred metres of Motomachi, but volcanic activity seemed to have switched to the southeast coast of the island, where plumes of discoloured water in the sea suggest submarine volcanic activity.

David Swinbanks

Biotechnology insured

Response to product liability

Washington

THE high cost of US product liability insurance has stimulated the Association of Biotechnology Companies (ABC) to plan a captive cooperative insurance company for the industry. If enough companies join the risk-sharing enterprise, it could offer "attractive" rates for insurance in respect of product liability as well as for the actions of directors and officers of companies, according to Bruce Mackler, ABC's general counsel.

Many smaller biotechnology companies have apparently not been able to obtain insurance at reasonable cost; some have had to proceed with none. ABC says one (unnamed) biotechnology company was recently quoted an annual premium of $400,000 for total coverage of $1 million.

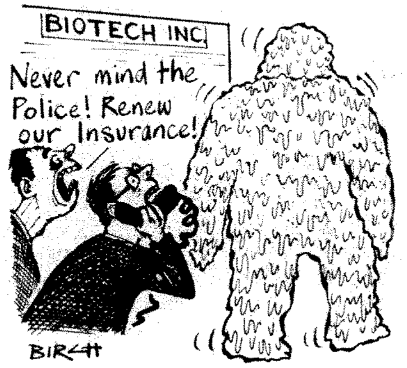
Part of the difficulty is the recent inflation of product liability claims, and of insurance premiums. The insurance industry is also under pressure from the number of product liability suits in federal courts, up by 750 per cent since 1974, and the average jury award, which has risen from $400,000 to more than $1.8 million during the same period, according to the American Insurance Association.

Insurers are especially fearful of the doctrine, recently introduced in California, of "joint and several liability", which means that a party only partly at fault can be obliged to become the "deep pocket" from which a whole claim must be met. Chemical, pharmaceutical, aircraft and car manufacturers are most at risk.

This, say insurance companies, is why they prefer to stick to the devils they know, such as automobile insurance and workers' compensation, than to take on unknown quantities such as biotechnology. They fear that huge claims may be made in the light of future knowledge, as they have been for medical products, when initial ignorance of harmful effects may not be a defence.

The ABC scheme was recommended by a study conducted by insurance brokers Johnson and Higgins of New York. Twenty-four companies from around the world, including well-known companies such as Celltech and Charles River, have chipped in enough to establish the biotechnology insurance cooperative legally in the state of Vermont; it should be ready for "activation" by early December. Membership will be open to any company in the field, but 30 members will be required to make it viable. The Industrial Biotechnology Association, which represents generally larger companies, has considered and rejected such a scheme, as many of its members are able to carry their own insurance.

The ABC captive insurer will provide only modest coverage at the outset: $1 million of either primary or "excess" coverage, with an aggregate ceiling of $3 million. But coverage would increase if the scheme expands; 44 companies have asked for quotes. ABC hopes the scheme will include insurance against environmental damage for such products as microbial pesticides, but admits that may be expensive. All participants will be inspected by risk management experts to ensure they are "doing things by the book".


BIOTECH INC
Never mind the Police! Renew our Insurance!
BIRCH

The American Insurance Association is generally sceptical of captive insurers, saying that insurance costs reflect the costs of meeting claims, no matter who does the underwriting. The insurance industry says that the provision of medical malpractice insurance through similar "risk retention groups" is now as expensive as through commercial agents.

Tim Beardsley

CNRS

Union power may not erode

The 20,000 staff, £600-million budget French national research council, CNRS, is to be given a lighter management structure and put back in the hands of the 'best' scientists after years of technocratic rule from the top and foot-dragging by trades unions.

That at least is the theory now being put out by the new CNRS administration, encouraged by the French prime minister, Jacques Chirac. Last week Chirac and his ministers gave approval for a new de-unionized structure for the research council's elected evaluation committee, the 1,000-strong Comité National. This gives back to CNRS power lost earlier in the year when the previous Comité National was declared illegal on a technicality, and clarifies at least part of the new government's policy towards science.

According to Serge Feneuille, director-general of CNRS, one of the new slogans will be that "science cannot be unionized", and the restructuring of the Comité National confirms it. In the elections, which will probably be complete by April 1987, at least half the candidates, especially the most senior, will no longer have to declare a union affiliation but may stand under their own name. Places on the committee for the engineers, technicians and administrators who are most unionized are to fall.

But although these developments are meant to loosen the hold of the unions over the system of promotions and grants, the new structure, according to its opponents, will shift power from the unions to another sector that could be equally conservative, university lecturers who do no research, but who will now be free to vote and stand for election. This move is a concession to university opinion that would have preferred CNRS to be dismantled. The resulting compromise, say critics, will merely hamstring the scientists in a new way. In the life sciences, many non-researching physicians from teaching hospitals will be joining the evaluation committees to make their opinions of French biology, and biologists, felt.

Whatever the politics of the Comité, there is some relief among 'les admissibles', the 512 young hopefuls who were appointed to full-time life-long CNRS research positions earlier this year by the previous Comité National, only to have the Comité declared illegal. At least now there is a legal Comité to look into their case, but, according to CNRS, they will have to take their examinations again. The same number of posts, 512, will be made available, and the resulting appointments will be counted as 1986 appointments with no bearing on the level of further recruitment in 1987. The proper 1987 appointments round (some 400–450 posts) will then take place in September and October, with a return to a normal spring timetable in 1988.

Meanwhile, Feneuille is known to have considerable interest in the introduction of short-term postdoctoral positions into the French research system which now locks itself into the appointment of a scientist for life at the age of 26. Feneuille believes that French industry, in which science is severely under-represented is in need of 'irrigation' with scientists. He argues that a system of postdoctoral appointments to CNRS in excess of the long-term need would allow the scientific irrigation of industry, as happens regularly in the United States and Britain. Such a change, however, could be political dynamite in France. **Robert Walgate**

Soviet academy

Marchuk in ministerial mode

The Soviet Academy of Sciences, already regarded by Soviet planners as the chief coordinator of all pure and applied research in the Soviet Union, is to acquire a new role. Addressing the Supreme Soviet (the Soviet Union's bicameral parliament) last week, Gurii Marchuk, the new president of the academy, said that it is to start producing forecasts for the major areas of science so as to identify the parts of the national economy ripe for development.

The new role of forecasting is part of a major reorganization of the structure and management system in line with the restructuring of the whole Soviet economy launched by Mr Mikhail Gorbachev last year. Under the new scheme, Marchuk said, the "specialized departments" of the academy will play a larger part, so as to direct scientific institutions towards the most important problems of fundamental and applied science.

How these changes will work is not yet clear. The major innovation so far of the Gorbachev era in science and technology, the "interbranch scientific-technical complexes" (MNTKs) are, in Marchuk's words "producing stress and difficulties". These organizations were supposed to "accelerate scientific and technical progress" by cutting across the interdepartmental bureaucracy and the barriers between science and industry. (At present the gap between initial breakthrough and large-scale production is said to be as much as 11 years.)

But as yet, the only MNTK to have received official commendation is the Paton Institute of Electric Welding of the Ukrainian Academy of Sciences, which has existed for many years and which served as the model for the MNTK concept.

Marchuk also repeated in his speech the targets outlined in the draft plan for 1987 (computers — both main-frame and personal; plasma, radiation and laser treatment of specialized materials; powder metallurgy; and biotechnology) and the Comecon Comprehensive Programme for Scientific and Technical Progress up to the year 2000. But, he warned, if the Soviet Union's "fine scientific and technical foundation for attaining the highest world standards" is to be properly employed, the planners, and in particular those involved in the capital construction of new enterprises, must play their part.

This is a new line. Hitherto, scientists have blamed industrialists for not using their results; more recently industrialists have begun to retaliate, saying that the results they were offered were unsuitable in practice. Marchuk now seems to be arguing that scientists can produce what industry needs only if the planners properly specify what is required.

But the real novelty of Marchuk's speech was not so much what he said but the fact that he addressed the Supreme Soviet in a virtually ministerial manner. As former head of the State Commission for Science, of course, he is no stranger to the Supreme Soviet. But the fact that he would have addressed it in this way at this early stage suggests that his presidency of the academy will be more vigorous than that of his predecessor. **Vera Rich**

Networking the begging-bowl

Strasbourg

The general assembly of the European Science Foundation (ESP), the most penniless and self-effacing of European science institutions, came to no particularly decisive conclusions last week. But the foundation did agree that it would ask its members (emphatically not governments but the research councils that governments support) for extra funds to keep its network programme alive.

That programme springs from an inquiry by ESF two years ago, at the behest of the European Commission, which showed that European researchers were too often isolated from each other. Largely because of a donation of a million French francs, the network programme last year brought 400 people together at 35 international meetings. The flavour, this year, is polar research, on which everybody is keen. But there are no funds, whence the begging bowl. Robert Walgate

Technology database

Best buys in UK research

EVERY researcher in Britain is to be mailed on behalf of an on-line computer database, designed to sell British talents to industry worldwide. Called British Expertise in Science and Technology (BEST), the commercially run database has been endorsed by the British government and the Science and Engineering Research Council (SERC).

Researchers in universities, polytechnics and the research councils are being asked to take part in the programme, in part intended to provide the supposed missing link between researchers and industrialists. Thirteen thousand records on researchers and their work now constitute the core of the database. A typical record will show a researcher's qualifications, publications, past and present research work, specialities, peripheral interests, industrial experience, current and previous positions. With the appropriate telecommunications link, the database can be tapped from any part of the world.

The database is part of the response by British universities to the publication in June 1983 of the report by the government's Advisory Council for Applied Research and Development (ACARD), which urged closer links on research between higher education and industry. In 1984, the universities set up a committee representing government departments and agencies as well universities and polytechnics. The committee selected the best proposal by Longman Cartermill, part of the Longman publishing house, from among other bids submitted.

Since the launch of the database in the spring, the operators claim to have attracted more than 100 industrial subscribers. Data entries to the system are free. Revenue will derive from subscribers, who are charged £200 for a single search or fees ranging up to £10,000 a year for full corporate use.

The terms of an agreement between SERC and the operators have yet to be concluded, but to start with, SERC will provide information on its research grants in return for access to the database for its own purposes. **Bill Johnstone**

- Meanwhile, the British Library is to launch its separate on-line computer database this week. That will focus more on research projects than researchers. The database contains 65,000 records of research being conducted in Britain. The system, called Current Research in Britain (CRIB), has evolved from an annual printed publication produced by the library. It has a £50-an-hour connection charge and is operated for the British Library by Pergamon InfoLine. □

Information technology

Filling the post-Alvey gap

A £1,000 million five-year plan to strengthen Britain's research in and application of information technology was proposed last week by the IT86 committee, set up this spring under Sir Austin Bide, the now-retired chairman of the Glaxo pharmaceutical group. The request for funds on this unprecedented scale is meant to continue the five-year Alvey programme of research and development in information technology, whose £350 million is now almost fully committed.

More than 300 projects have been approved under the Alvey programme, but the Bide committee considers it too soon to evaluate their success except to say that they have been successful in drawing industry and universities together more closely "than was previously the case". The committee also says the programme has stimulated collaboration between companies normally in competition.

Of the total £1,000 million, the committee proposes that roughly £425 million should come from the British government and the remainder from industry. The government contributed £200 million to the total cost of Alvey projects.

The Bide committee estimates that the research element in the continuation programme will cost £550 million, of which the government is expected to contribute £300 million. As under Alvey, academic research would be fully supported by government, but research by industrial companies receives only half-support.

The committee also endorses some of the conclusions of a report by the British government's Advisory Council for Applied Research and Development (ACARD) which, earlier this year, was critical of the British software industry. That document forecast a continued increase in the British trade deficit in information technology.

Much of the Bide committee's emphasis is on the need for converting research into marketable products. The committee says that applications schemes must involve collaboration between researchers and industry and that academic contributions to joint ventures should "go well beyond" traditional information-technology departments. The report also includes a long shopping-list of projects to be supported. Among applications projects, it singles out security systems, clinical data processing, interactive distance learning aids, distance learning for management, manufacturing control, electronic funds transfer at point of sale and software for railway signalling.

On research, the Bide committee recommends a number of areas to which effort should be directed, including systems design and management and new programming tools. Not all the support now requested need be new money; part of it may come from the funds already earmarked for the second phase of the European Strategic Programme for Research and Development in Information Technology (ESPRIT).

No doubt in recognition of the difficulty of extracting large amounts of money from the budgets of individual government departments, the committee says that the funds should come not only from the original sponsors of the Alvey programme (the Departments of Trade and Industry and of Defence and the Science and Engineering Research Council) but from other departments with an interest in the applications suggested, either as sponsors or as users. **Bill Johnstone**

The Straits of Gibraltar, with Spain to the left and North Africa to the right. On page 313 of this issue we publish a 20-page supplement on science in the Iberian peninsula. Photo: Oblique view from Skylab, courtesy NASA.

Dissent disallowed in Singapore

SIR—There exists an interesting juxtaposition of events in Singapore this week. On the one hand, scientists from all over the world are gathering for the Fourth Federation of Asian and Oceania Biochemists (FAOB) meeting, to present and discuss work freely. This meeting is important, as it is timely, to the powers that be in Singapore, in that it precedes the opening of the new Institute for Molecular and Cell Biology there, for which the talents of the international scientific community are being anxiously courted. In stark contrast, the voice of dissent in Singapore's parliament, Mr Joshua Jeyaretnam, is effectively silenced this week; together with the chairman of the opposition party, Mr Wong Hong Toy, he finds himself in the cells of Singapore's Changi Prison.

The history of events that have led to fines, imprisonment and the removal of the leader of the opposition from his seat in parliament is adequately dealt with in such publications as *Time Magazine* (8 September 1986), *Financial Times* (12 November 1986), *Far Eastern Economic Review* (4 September and 20 November 1986) and the *Asian Wall Street Journal*. As a result of these articles, in which the motives of the government and the independence of the judiciary are called into question, two of these publications have incurred sanctions. The net overall effect of these events is to suggest that the voice of dissent in Singapore, be it from individuals, editors or members of parliament, is to be silenced at any cost. It is against this political backdrop that the Singapore government, advised by a number of eminent scientists from the United Kingdom, the United States of America and West Germany, seeks aggressively to recruit internationally scientists upon whose cooperation the success of the FAOB meeting and that of the Institute for Molecular and Cell Biology rests.

This situation, taken as a whole, raises some quite difficult questions, addressed especially to those advising or considering working at the Institute for Molecular and Cell Biology in Singapore. When a scientist enters into a commitment of research in Singapore, does he retain his right to comment on what he sees? If the independence of the judiciary and the freedom of the press are called into question, who is left to take the responsibility of safeguarding the rights of the individual to criticize? This responsibility sometimes comes to rest in unexpected quarters. In this case, because their services are of value to the Singapore government, some of this responsibility lies squarely with the international scientific community. It is up to us to satisfy ourselves (the facts are on record) that the individual in Singapore who wishes to voice legitimate criticism has the freedom to do so. If this freedom does not prevail, then the scientific community at large, being in a position of influence, must not flinch from the responsibilities associated with this position, and must use this influence to foster the necessary change of attitude within the government in Singapore.

No government should assume that it can, through institutions such as the Institute for Molecular and Cell Biology in Singapore, exploit the talents of the international scientific community and expect the scientists concerned to turn a blind eye to the fate of the people in whose country they work.

LOUIS MAHADEVAN
London NW6, UK

Fluoridation

SIR—The 1983 Anglesey study[1] does not demonstrate the effectiveness of fluoridation, as Francis B. Reed claims[2]. Apart from its questionable late selection of a known high-caries non-fluoridated control population different from that originally planned, without pre-fluoridation information on the populations being compared — making its "strictly blind conditions"[2] worthless — its concluding assertion of "the universal finding that caries experience in a fluoridated community is consistently lower than in neighbouring non-floridated communities" is shown to be untrue by the many studies cited in Diesendorf's Commentary[3]. In New Zealand we have dental health figures for our entire child population, not just for samples as in "fluoridation trials". The figures for Auckland show that the children in the non-fluoridated part of the city require fewer fillings, not more[4]. The non-fluoridated off-shore island children also need fewer fillings, just like the fluoridated island of Anglesey.

JOHN COLQUHOUN
Education Department,
University of Auckland,
Private Bag, Auckland,
New Zealand

1. Jackson, D. *et al. Br. Dent. J.* **138,** 165–171 (1975); **158,** 45–49 (1985).
2. Reed. F.B. *Nature* **323,** 198 (1986).
3. Diesendorf, M. *Nature* **322,** 125–129 (1986).
4. Colquhoun, J. *Community Dent. Oral Epidemiol.* **13,** 37–41 (1985).

UK performance

SIR—Your contributors have recently reviewed two influential reports on British performance in research[1,2]. One concludes that "the main findings are that in basic research in science as a whole, UK performance has deteriorated since the early 1970s both in absolute terms and relative to other countries, notably FRG and Japan". The figures that most strikingly emphasize this point are the citation rates.

This broadbrush conclusion disguises a more precise relationship. The relative proportions of each of the competitors' total research budgets spent on defence in the early 1980s have respectively been 0.5, 0.2 and 0.02, demonstrating a clear negative correlation between the proportion of research funding for defence and the corresponding civil research performance.

One can only guess why this should be so, if indeed the relationship is real at all, let alone representative of the world at large, derived as it is from only three (very important) examples. Yet if it is real it smacks of a sort of logic: the greater the proportion spent on defence, the greater the number of defence scientists required to spend it and consequently, from a limited resource, the smaller the number in civil research; also the greater the proportion spent on defence, the greater is the restriction on research reports available for citation. The former points towards a real problem for UK civil research; the latter to a possible mitigating explanation of its apparent recent lacklustre performance.

BRUCE DENNESS
Bureau of Applied Sciences,
Wydcombe Manor,
Whitwell,
Isle of Wight PO38 2NY, UK

1. *Evaluation of National Performance in Basic Research*, ABRC Science Policy Studies No. 1 (Advisory Board for the Research Councils, London, 1986).
2. Martin, B.R., Irvine, J. & Minchin, N. *An International Comparison of Government Funding of Academic and Academically Related Research* – Vols 1 & 2, ABRC Science Policy Studies No. 2 (Advisory Board for the Research Councils, London, 1986).

Backfiring boycott

SIR—The signatories of the Strategic Defense Initiative (SDI) boycott pledge (see *Nature* **323,** 747; 1986) may find themselves in an embarrassing position, should the United States and the Soviet Union agree on how to regulate laboratory research on SDI. Would they recant their position, and on what grounds? The research programme (and that is all that the academics are involved in) will certainly not become technically any less dubious simply by receiving Mr Gorbachev's blessing. And were the Western scientists to continue their boycott unilaterally, they would simply be paving the way to a technological breakthrough by the Soviets, tempting them to break out from the Anti-Ballistic Missile Treaty. Is this course of action any less misguided than participation in SDI research? The most disturbing feature of the anti-SDI pledges circulating in the United States and now in Europe is that they are black-and-white statements that do not allow for any shades of grey.

NICHOLAS ZUMBULYADIS
686 Seneca Parkway,
Rochester, New York 14613, USA

How to make antimatter last

An experiment at CERN has shown that antiprotons can be kept for several minutes in an electromagnetic trap. But the prospect of making reality of the science fiction remains remote.

SCIENCE fiction writers, ingenious people as they are, include among the standard props of their trade a ready supply of antimatter, the most portable of all high-explosives. During the past few years, there has also been a disturbing trickle of accounts in the popular press suggesting that the world's accelerator laboratories are hard at work on the development of techniques for making antimatter for use in military operations. A glance at one of the first published accounts of success in storing antiprotons (*Phys. Rev. Lett.* **57**, 2504: 1986) should persuade all those concerned that it will be some time before the fiction can become reality.

For other reasons, the new development is interesting and important. Working at CERN, the European high-energy physics laboratory at Geneva, a group from the University of Washington at Seattle (Gabrielse, Kei, Helmerson, Rolston, Tjoekler and Trainor), two people from the University of Mainz (Kalinowsky and Haas) and one from the Fermi National Accelerator Laboratory (Wells) have been able to decelerate antiprotons to the point at which their energy is sufficiently small to allow them to be captured in an electromagnetic trap. The experiment is a technical achievement in its own right, but the prospect that it will now be possible to study antiprotons essentially at rest, and stable for some minutes, raises the prospect of a series of previously impossible experiments.

CERN's advantage in this connection is the facility called the Low Energy Antiproton Ring (LEAR), which is a device for storing bunches of antiprotons produced by the Super Proton Synchrotron (as modified so that protons and antiprotons travel in opposite directions) and reducing their individual energy by four orders of magnitude to 5 MeV or thereabouts. Novel experiments are possible even at this energy (see Batty, C.J., *News & Views, Nature* **323**, 487; 1986). But even 5 million volts is far greater than the electrostatic potential well of the kinds of traps that might be constructed in the laboratory.

The obvious way to make possible the trapping of antiprotons would be a further elaboration of the LEAR storage ring, but although CERN has ambitions in that direction, the extra equipment has not yet been built. Accordingly, the Washington–Mainz group has fallen back on the simplest of all devices for robbing fast charged particles of energy — that of making them traverse a dense target of matter, a solid beryllium window. In practice, the experiment uses virtually instantaneous bunches of antiprotons, each of them containing between 10 million and 1,000 million particles.

Inevitably, the process is inefficient. With the thickness of the window chosen to be the range of antiprotons, at the energy discharged from LEAR, in beryllium, only about 1 in 10,000 of the original bunch emerges from the target with an energy less than 3 keV, which means that typical bunches leave only some 10,000 antiprotons for catching in a trap.

That consists of two components, of which the simplest is a powerful axial magnetic field to prevent slow antiprotons moving too quickly away from the axis of the experiment. The other is a conceptually simple but technically sophisticated pair of electrodes, each with a hole in the centre (a third of a centimetre across) to allow the passage of the antiproton beam, and roughly a dozen centimetres apart. Before the arrival of an antiproton bunch, the first electrode is at earth potential and the second, or more distant, has a potential of −3 kV so as to repel antiprotons less energetic than 3 keV. But while the bunch is travelling the length of the trap, the voltage of the first electrode is switched, in a matter of nanoseconds, to match the potential of the first.

For the time being, the objective is to see how efficient such a simple trapping technique may be. As these things go, the mechanism seems to work as well as anybody expected in advance. The population of antiprotons can be measured at any time by abolishing the repulsive potential on the more distant of the two electrodes, whereupon the antiprotons are guided by the axial magnetic field to a target where they annihilate, yielding pions which can be counted. Apparently, the counting equipment is not yet fast enough to measure the population of a newly filled trap, but typical records show that, after a millisecond, the trap may contain 300 or so antiprotons. One of the published records shows that after storage for 100 seconds, it was still possible to record 31 antiprotons, spread over a mere 5 microseconds. On balance, it seems as if it is possible to catch roughly 3 in a million of all the particles discharged as a bunch from LEAR.

At this rate, sheer numbers are less important than the other properties of the trap, which will determine what use can be made of a supply of essentially static antiprotons. The authors of the experiment brim over with suggestions for what may happen next. One scheme is to make single atoms of anti-hydrogen, perhaps by building two traps alongside each other, one for antiprotons and one for positrons. Firing a beam of positronium atoms (pairs of positive and negative electrons) at antiprotons in a trap might be a simpler way of accomplishing the same goal. The prospect of precision measurements of quantities such as the antiproton mass, or even of the gravitational acceleration of antiprotons, will intrigue many observers of this new development.

Technically, the improvement of the present means of trapping antiprotons is an obvious line of evolution, but the authors write of the still more promising line of development in which antiprotons trapped as they describe would be transferred, bunch by bunch, to a trap whose geometry is not constrained by the need to capture as many particles as possible from LEAR. Such a device would therefore be better suited for keeping even these unstable particles for long periods of time, a day or even longer. That is what will have to be done if some of the more interesting measurements now suggested are to be carried out.

All this is of course a long way from what the science fiction writers require of antimatter. While the best that people can do is to store a few hundred antiprotons for an hour or so, the prospect of antimatters weapons will remain very distant. The essential difficulty is a score or so of orders of magnitude. True, the annihilation of an antiproton by, say, a proton of the kind that turns up in ordinary matter is more energetic, by a factor of say a hundred, than the fission of a uranium nucleus. But Avogadro's number (6.10^{23}) works emphatically in the other direction. So long as numbers of antiprotons that people can collect in antiproton traps are small enough that it is feasible to think of counting them, the popular newspapers' accounts of what high-energy physics is really for will remain in the realm of fantasy. It is nevertheless outrageous that there have been in the past few years, such a torrent of tales along these lines. They will always be fairy stories. So, too, perhaps sadly, will be the relevant parts of science fiction.

John Maddox

Evolution

Contemplating life without sex

John Maynard Smith

FROM the selfish-genetic point of view, there are obvious advantages to parthenogenesis. A gene that suppressed meiosis in female diploids would be transmitted to all female offspring instead of only to half and thus would, all else being equal, double in frequency in each generation. Why then, does parthenogenesis not replace sex? The answer has to be that all else is not equal. Why not? One good reason for studying parthenogens is for the light they may shed on the evolution of sex, and the problems of parthenogens were the topic of a recent symposium* from which it is clear that for one reason or another it is very difficult to give up sex once you have it.

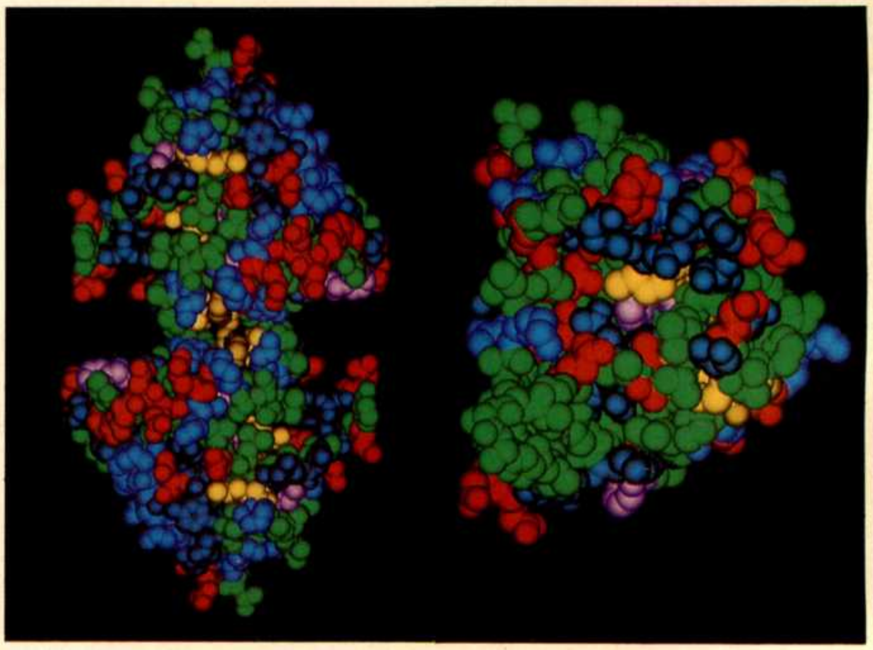
Most existing parthenogens have evolved from sexual forms and are hampered to varying degrees by the persistence of mechanisms appropriate only to their ancestral lifestyle. As so often, the fruitfly *Drosophila* provides a good example. Females of many species lay an occasional egg that develops without fertilization. But it is a long way from the occasional egg to functional parthenogenesis (A. Templeton, Washington University, St Louis). Parthenogenetic laboratory strains now exist in several species, including *D. melanogaster*, but it took much selection to establish them. Before selection, unmated females lay few eggs, and even fewer start development. Of those that do start, most die because of genetic homozygosity resulting directly from the mechanism of parthenogenesis. The mechanism is known as automixis: the egg undergoes meiosis, and diploidy is restored by the fusion of two of the four products. Because natural populations carry many deleterious recessive genes, automixis results in heavy mortality. The only species which is parthenogenetic in nature, *D. mangabeiri*, has overcome this problem in an ingenious way. First, the two haploid nuclei that fuse are those that are separated at the first meiotic division; second, there is no crossing over. Together these devices ensure that the offspring are genetically identical to the parent: heterozygosity is preserved. Crossing over is prevented by inversions.

The picture that emerges from these studies of *Drosophila* is probably typical. The disadvantages suffered by recently arisen parthenogens are the direct result of their sexual ancestry; but these 'sexual hang-ups' can be overcome by selection. The precise means of overcoming them varies. Some parthenogenetic animals, and many plants, opt for abandoning meiosis altogether — apomixis; this is the pattern in the cyclically parthenogenetic aphids and cladocerans. Alternatively, in the parthenogenetic lizards of the genus *Cnemidophorus*, there is a premeiotic doubling of the chromosomes, followed by the pairing of sister chromosomes in the first meiotic division. This has the same genetic consequence as the meiosis of *D. mangabeiri*: offspring are genetically identical to their parents, and do not suffer the enforced genetic homozygosity that is the usual result of automictic parthenogenesis. The parthenogenetic species *C. uniparens*, which is a highly heterozygous species hybrid, consists of a single clone: members of the species accept scale grafts from each other.

Another common type of sexual hang-up in parthenogenesis is the persistence of the need for sperm to trigger development — pseudogamy (that is, the stimulation of egg development by a sperm whose chromosomes do not participate in forming the new individual).

In some lower vertebrates (discussed by H. Hotz and T. Uzell, University of Illinois and R.C. Vrijenhoek, Rutgers University), the sperm is probably needed to provide a centriole. (No such need, however, exists in lizards.) Pseudogamy is not uncommon in insects. In plants, a nucleus from the pollen tube may be needed to fertilize the endosperm, even if the embryo is parthenogenetic. Some parthenogenetic frogs (*Rana*) and fishes (*Poeciliopsis*) have evolved what is perhaps the strangest mechanism of all for dealing with this requirement. It is known as hybridogenesis. *Rana esculenta*, for example, is a hybrid between two sexual species, *R. ridibunda* and *R. lessonae*. In the germ line, *lessonae* chromosomes are thrown out, so that the egg contains only the *ridibunda* component. If, as is typical, the hybrid female mates with a *lessonae* male, the hybrid *R. esculenta* is reconstituted. Clearly, pseudogamous and hybridogenetic forms cannot wholly replace their sexual ancestors. In some major taxa, for example mammals and gymnosperms, parthenogenesis has not been reliably recorded, but the nature of the hang-up is not known.

Despite the widespread occurrence of sexual hang-ups there are many parthenogens that seem to have escaped the disadvantages derived from their sexual ancestors. The taxonomic distribution of parthenogens, however, suggests that they may be successful in the short run,



THESE three-dimensional views of β-lactoglobulin show two identical monomers tenuously linked. Each monomer is similar to plasma retinol-binding protein and each contains a non-polar pocket similar to that occupied by retinol in retinol-binding protein. The non-polar pocket is in the centre of view in the right-hand figure. The function of β-lactoglobulin, a milk protein, has long been unknown; these new models suggest that it is a transport protein. The structure is described in more detail by M.Z. Papiz and colleagues from the Universities of Leeds, Edinburgh and Uppsala, on pages 383–385 of this issue. The colours represent different types of amino-acid residue: red, acidic; dark blue, basic; light blue, polar; purple, aromatic; yellow, sulphur-containing; green, non-polar. Computer program by Dr H. Nakamura. □

---

*A meeting on parthenogenesis organized by L. Kirkendall and N.C. Stenseth at Finse, Norway, 17–22 September 1986.

but are short-lived on an evolutionary timescale Thus, most parthenogens have close sexual relatives It is rare for a taxon higher than a species to consist wholly of obligate parthenogens The parthenogenetic order Bdelloidea of rotifers, and the tribe Tramini of aphids, are important exceptions that challenge further research

Among higher plants, taxa such as *Taraxacum* (dandelions) and *Rubus* (blackberries) that contain many parthenogens also contain sexual species, and new clones are continuously arising Two theoretical reasons have been suggested for this short evolutionary life expectation First, evolution to meet changed circumstances will be slow in the absence of genetic recombination Second, there may be a tendency to accumulate slightly deleterious mutations (Muller's ratchet[1]) Although the theory is clear, it is hard to provide empirical support A major difficulty lies in deciding how old particular clones may be, and in distinguishing between variation that has arisen by evolution within a clone and that which reflects multiple origins from sexual ancestors No-one was prepared to make even a rough estimate of the age of any clones

The simplest picture is as follows Most parthenogenetic 'mutants' fail to establish themselves because of various sexual hang-ups Some, under intense selection, overcome these difficulties and achieve short-term ecological success Despite their twofold reproductive advantage, however, success is short-lived on an evolutionary timescale because of their relative inability to adapt, and perhaps because of the accumulation of mutations

Is any further explanation needed? Williams[2] argued that it is, essentially because the coexistence of closely related sexual and asexual populations implies that sex must confer some immediate advantage He thinks that this advantage arose from the genetic variability associated with sexual reproduction This can take two forms first, a sexual population will be more variable than an asexual one, and second, the offspring of a single sexual female will vary more than the offspring of an asexual one

The variability of a sexual population may be relevant in the following way A single clone, with a single genotype, is unlikely to displace its sexual ancestor over its whole ecological and geographical range Vrijenhoek has called this the frozen niche hypothesis (Bell's[3] tangled bank model is, I think, similar), and argues that, as the number of different clones increases, the relative success of the sexual species should decline He finds that, in *Poeciliopsis,* clones do differ physiologically and in their ecological adaptations, and that the proportion of the population that is sexual declines as the number of distinct clones in a stream decreases In *Poeciliopsis*, however, parthenogens need to be mated, so they cannot wholly replace their sexual ancestors Clonal diversity is now being studied, using biochemical markers (and, in vertebrates, immunological methods), in many plants and animals, and this should help us to decide on the importance of population variability in conferring an immediate advantage on sexual populations

The Williams 'balance' argument may also be relevant to cyclical parthenogens such as cladocerans (M Lynch, University of Illinois) and aphids (N Moran, University of Arizona) We must be cautious, however, because there is often an ecological difference between the sexually and asexually produced eggs For example, in *Daphnia* the resistant winter eggs are often produced sexually If this was a universal rule, the maintenance of sex could be explained by the need to produce winter eggs But there are obligately parthenogenetic strains of *Daphnia,* some of which produce winter eggs asexually The situation is similar but more complex in aphids Most aphids have two plant hosts, and the sexual forms attack the secondary host No obligately parthenogenetic aphid has retained the ability to utilize two hosts, but there are many cyclically parthenogenetic aphids that are confined to a single host, and these have given rise to many obligately parthenogenetic strains Some modification of the frozen niche hypothesis, therefore, may be needed to explain why all cladocerans, and aphids with only one host, do not wholly abandon sex

Williams argues that genetic variation among the offspring of a single female may also be important He uses the famous analogy that a parthenogenetic female is like a man who buys 100 tickets in a lottery and finds that they all have the same number I have argued[4] that this model only works if the offspring of one female compete with one another, but showed that, provided that there is sib competition, the model can give rise to a twofold advantage to sex

But do we need to take into account the genetic variability of individual families, or are the other explanations outlined above sufficient to explain the failure of parthenogens to replace sexual forms? There is one context in which within-family variability may be relevant

Consider a population of diploid parthenogenetic females, that is, females that produce some of their offspring parthenogenetically and some sexually Why does such a population not evolve towards obligate parthenogenesis? From the point of view of its own proliferation, there would be an intrinsic advantage to a gene that suppressed meiosis, and it is hard to believe that there would not be some genetic variance affecting the proportion of offspring produced asexually and on which natural selection could work We cannot appeal to the greater variability of the sexual population, because the range of genetic variability is the same in the sexually and asexually produced offspring The sib competition model might, therefore, be necessary

Do such facultative parthenogens exist? In animals, I know of no clear example Populations with a female-biased sex ratio usually consist of a mixture of obligate parthenogens and a bisexual population But there are many case that merit further investigation In plants, there do seem to be examples One, the diploid form of *Potentilla argentea,* was described (S Asker, Lund) These cases need genetic and ecological study

I formed the strong impression that the study of parthenogenesis is, at last, beginning to tell us something about the evolutionary significance of sex Much remains to be done The Bdelloids remain something of an evolutionary scandal We need to know the age of at least some natural clones, the extent of genetic diversity that can arise within them, and whether deleterious mutations accumulate in them by the ratchet mechanism Further studies of clonal diversity will show how important the frozen niche and tangled bank models in fact are The suggestion[5,6] that sex is needed in the evolutionary race against parasites should be investigated The few clear cases of diploid facultative parthenogenesis should tell us whether sib competition is relevant I think that the problem of the maintenance of sex in higher eukaryotes is different both from the problem of the evolution of recombination in prokaryotes and from the origin of meiosis and syngamy in single-celled eukaryotes The three problems may have different solutions Indeed, I am reasonably sure that the first two do have different solutions, because the primary function of recombination in prokaryotes is probably the repair of damaged DNA Although DNA repair is also needed in eukaryotes, I do not think that the suggestion[7] that meiosis and syngamy serve a repair function is tenable indeed, one of the reasons why parthenogens are interesting is that they demonstrate that this suggestion is mistaken As far as the maintenance of sex in organisms like ourselves is concerned, I think the main theoretical explanations have already been formulated, and our task is to decide on their relative importance □

1 Muller H J , *Mutat Res* 1, 2 (1964)
2 Williams, G C *Sex and Evolution* (Princeton University Press 1975)
3 Bell, G *The Masterpiece of Nature* (University of California Press, Berkeley, 1982)
4 Maynard Smith, J *J theor Biol* 63, 245 (1976)
5 Jaenike J *Evol Theory* 3, 191 (1978)
6 Hamilton W D *Oikos* 35, 282 (1980)
7 Bernstein H Byerly H C Hopf F A & Michod R E *Science* 229, 1277 (1985)

John Maynard Smith is Professor of Biology at the University of Sussex, Falmer, Brighton BN1 9QG, UK

Materials science

Nuts to tough ceramics

Paul Calvert

MACADAMIA or Queensland nuts are much used in high-class cookies but are very difficult to remove from their shells without special equipment In a recent article, J S Jennings and N H Macmillan (*J Mater Sci* **21**, 1517, 1986) discuss the properties of the fibrous macadamia nutshell and shed some light on the prospects for fibre-reinforced ceramic materials which are currently thought to be the route to tough ceramics for use in engine components

Jennings and Macmillan looked at the microstructure of the nutshells and measured the mechanical properties of various sections of the shells in the wet and dry states It is very difficult to obtain reproducible mechanical measurements on biological materials because the properties are frequently dependent on the age and moisture content of the tissue and vary with orientation and position in the organism as well as being vulnerable to flaws This creates something of a philosophical problem for engineers who are faced with a natural system, when they are more used to uniform properties measured on large pieces of material In this (nut) case the properties are surprisingly uniform and are best described as those of 'isotropic wood' That is, the moduli, strength and toughness are comparable with the geometric averages of those parallel to the grain and across the grain of a softwood A nice recent example of this grain effect is a study of kelp (Vincent, J F V and Gravell, K *J Mater Sci Lett* **5**, 353, 1986), which forms 'leaves' by the tearing action of the sea along the weak direction parallel to the fibres The work of fracture decreases by 500 times from fracture across the fibres to fracture parallel to the fibres

The Young's modulus (elasticity in tension) of the nutshells is about 5 GPa, where wood is about 10 GPa parallel to the grain and 0 5 GPa across the grain The (fracture) strength of the shells is 50–80 MPa, where wood is about 100 MPa along the grain and 5 MPa across In terms of a complete nut this corresponds to a hungry gourmet needing to exert about 50 kilos of force on a normal pair of nutcrackers The fracture toughness is calculated from the stress needed to propagate a crack in a notched sample, and is thus a measure of one's ability to open the nut by making a small cut and levering on it This value is in the range from 0 8–2 MPa m$^{1/2}$ for the nutshell whereas wood is 1 and 2 MPa m$^{1/2}$ parallel and perpendicular to the grain, respectively As Jennings and Macmillan point out, these properties are not terrific when one takes into account the fact that the density of nutshells is 2 5–3 times that of wood The advantage of this design seems to lie in the isotropic properties which ensure that the shell is not weak along any particular direction and so cannot be split like wood, on the same principle that plywood gains strength by having layers with crossed grains This toughness precludes attacking the nut by a shrewd wallop with a hammer and chisel

The shell is made up of fibrous cells, each about 30 µm in diameter, which are arranged in bundles of tens or hundreds of parallel rods The bundles then pack to fill the volume When the shell breaks, the fracture either goes through the rods or around them, depending on whether the crack is perpendicular or parallel to the rods The toughness then arises from the need for the crack to change direction as it traverses the shell, possibly with some microcracking within and around the fibres that significantly adds to the energy needed to drive the crack This is illustrated by recent work of A R Sanadi, S V Prasad and P K Rohatgi (*J Mater Sci Lett* **5**, 395, 1986), who looked at fracture of polyester resins reinforced by fibres of sunhemp They observed extensive splitting and deformation of the fibres during fracture of the composite, and argue that this would give higher toughness than found in an equivalent composite reinforced with a brittle fibre such as glass The sunhemp composite has 15 times the work of fracture of the polyester resin alone

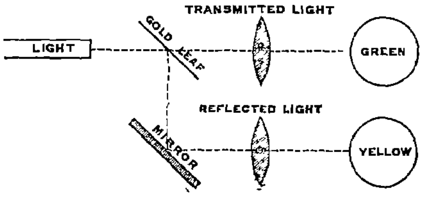
At a recent meeting on ceramics in Pennsylvania, United States (15–19 June 1986), A G Evans (University of California, Berkeley) summarized the requirements for the successful application of ceramic (for example, alumina or silicon nitride) parts in engines These were a toughness of 20–25 MPa m$^{1/2}$, a strength of 300 MPa and a creep strain of less than 1 per cent at high temperature (>1,000°C) All these properties can be achieved, but not all in the same material at once and not reliably The toughness of the ceramic can be raised by two methods one is to incorporate some tetragonal zirconia which raises the toughness by a martensitic transformation process similar to that in steel from 3–5 to 10–20 MPa m$^{1/2}$ This has the disadvantage that the toughening is not stable to high-temperature treatments The second is to incorporate fibres or whiskers of another ceramic to make a ceramic-toughened ceramic composite Because we want to make moulded parts rather than bars or sheets, the fibres must be randomly oriented to produce a structure very similar to that of the macadamia shell This structure is quite different from conventional composites where the fibres are either all parallel or are arranged in layers and where a hard, strong fibre is used in a soft but tough matrix, as in glass-fibre reinforced plastics

Silicon carbide whiskers about 0 6 µm long are commercially available from pyrolysed rice hulls These whiskers have been used to reinforce alumina, producing a doubling in toughness to 6–9 MPa m$^{1/2}$ Because the system was hot-pressed, the fibres were in a random-in-plane orientation rather than isotropic At loadings higher than 20 vol per cent, the whiskers did not increase toughness because they cause extensive cracking in the alumina during sintering (Wei, G C and Becher, P F *Am Ceram Soc Bull* **64**, 298, 1985) The questions that we now wish to ask the nut are what is the maximum toughening we can expect from whiskers in hard matrix systems, and how can we achieve high loadings of fibres into an isotropic, packed-bundle arrangement? □

Paul Calvert is in the School of Chemistry and Molecular Sciences, University of Sussex, Brighton BN1 9QJ, UK

100 years ago

WE are apt to think of gold as being essentially and distinctively golden-yellow, it may, however, possess a wide range of colours A leaf of gold when seen by transmitted light is either green or blue, according to its thickness Here is such a leaf of green gold as seen when light is actually sent through it, so as to project a green disk on screen A portion of the light will be *reflected* from its surface, and this reflected ray may be caught in a mirror and thrown on the screen so that you have, shown side by side, the green disk of transmitted light and the golden *one of reflected light from the same leaf of gold*



It would be easy to show that light is similarly affected by other metals, but I have selected gold for the purpose of illustration because it is easy to maintain it in a state of purity, however finely divided it may be

We must therefore modify any views we may have formed as to a metal having exclusively a special colour of its own, because it will be evident that a particular colour is only due to a definite state of arrangement of its particles

From *Nature* **35**, 109, 2 December 1886

Semiconductors

An optical Stark effect observed in gallium arsenide

John Ryan

Electromagnetic radiation interacts with atoms to modify many of their fundamental properties including the atomic energy levels themselves The most famous example of this is the Lamb effect, which occurs spontaneously in the absence of an applied electromagnetic field (in other words the field is 'virtual', arising from quantum fluctuations in the vacuum), in atomic hydrogen small energy-level shifts of the order of 1 GHz arise from the effect In condensed matter, the intrinsic energy levels are broadened considerably by interatomic interactions so that such small shifts are not normally detectable However, André Mysyrowicz and colleagues at École Polytechnique in Paris now report (*Phys Rev Lett* **56**, 2741, 1986) the observation of a remarkably large shift of around 500 GHz in the energy levels of the semiconductor gallium arsenide, GaAs On the basis of their results, the authors believe that the effect can be observed quite generally in semiconductors, and that it has potentially important applications in high-speed optical switching

The semiconductor studied by the Paris group is a composite material known as a quantum-well structure it consists of thin slabs of GaAs containing about 50 atomic layers sandwiched between equally thin slabs of GaAlAs such that the potential barrier formed at the interface between the two materials causes electrons and holes to be confined to the GaAs The energy state under investigation was the intrinsic exciton, which is a hydrogen-like bound state of an electron and hole having a Bohr radius of approximately 100 Å In GaAs quantum wells, the resonant generation of excitons by light gives rise to a well-defined absorption line which is sufficiently sharp to be a useful probe of electric-field-induced energy shifts

In the experiment, an electromagnetic field was applied in the form of an intense ultra-short laser pulse of 150 femtosecond duration The electric field produced is equivalent to a potential difference of about a tenth of a volt over the exciton, which is in fact greater than the electrostatic field experienced by the electron or hole due to the presence of the other Under these conditions large energy shifts are expected

Previous experiments by the authors (Hulin, D *et al Phys Rev* **B33**, 4389, 1986) and by a group at AT&T Bell Laboratories (Schmitt-Rink, S , Chemla, D S and Miller, D A B *Phys Rev* **B32**, 8027, 1985) had previously detected a pronounced blueshift of the exciton absorption line under conditions of resonant pumping, that is when the laser wavelength coincides with the absorption line This effect was shown to be due to exciton–exciton interactions, with the shift persisting during the exciton lifetime, typically several nanoseconds In its latest experiment, the Paris group now detects a large blueshift for non-resonant below-band-gap pumping (where no excitons are generated) which persists only during the pump laser pulse Mysyrowicz interprets this as evidence of an electric-field-induced energy shift or optical Stark effect

To explain the result, the authors adopt a quantum-electrodynamical model of 'dressed' excitons which are mixed states of excitons and photons The state with n_e excitons and n_p photons $|n_e, n_p>$ is nearly degenerate with the state $|n_e+1, n_p-1>$, and when the exciton–radiation interaction is included, the new mixed states are found to be split in energy by an amount which depends on the strength of the field This model therefore explains qualitatively the observed blueshift of the exciton absorption line, the magnitude of the shift can also be predicted theoretically, but there is some uncertainty in the actual value of the electric field strength inside the semiconductor that prevents a precise quantitative comparison

The Paris group has already extended its measurements to bulk GaAs and again detected an optical Stark effect These are important fundamental observations that open the way to the study of quantum-electrodynamical effects in condensed matter For the time being, however, the suggestion of practical applications looks somewhat premature because of the very large field strengths required ☐

John Ryan is at the Clarendon Laboratory, Oxford OX1 3PU, UK

Plant pathology

Host antifungal agents

Richard A Dixon

Unlike animals, which can defend themselves against microbial attack by producing circulating antibody proteins, plants respond to fungal and bacterial pathogens by synthesizing a number of proteins that produce a localized 'antimicrobial environment' around the invading pathogen[1,2] Lectins are thought to play a role as they have the potential to bind fungal or bacterial cell walls[3], and hydrolases (chitinase and 1,3-β-D-glucanase) have the potential to degrade macromolecular components of fungal walls The report by Schlumbaum *et al*[4] on page 365 of this issue now suggests that chitinase is an important component of the plant's antifungal arsenal and shows that the antifungal activity of plant lectins results from contamination with chitinase

The microbially induced plant gene products that have received most attention have been the enzymes involved in the synthesis of low molecular mass antifungal products termed phytoalexins[1,2] A large body of mainly correlative evidence suggests that the *de novo* synthesis of phytoalexins from intermediary metabolites, a process which in legumes such as bean probably requires the induced transcription of about 10 different genes, is a major determinant of active defence against microbial pathogens This complexity is daunting for the development of disease-resistant plants using gene-transfer techniques, hence the attraction of chitinase which is a single antifungal product

Previous reports have shown that this enzyme has no apparent endogenous substrate in plants but can release chitin oligomers from purified fungal cell walls, and is induced in plant tissues by fungal infection or exposure to the gaseous plant growth regulator ethylene Schlumbaum *et al*[4] now show that purified bean chitinase is highly inhibitory to the growth of the fungus *Trichoderma viride,* at concentrations similar to those found in crude extracts from infected or ethylene-treated bean leaves The antifungal activity of bean leaf extracts is almost completely inhibited by anti-chitinase antibody, suggesting that chitinase is the major protein inhibiting the growth of *Trichoderma* that is present in these extracts At similar concentrations, chitinase shows the same antifungal activity as the bean phytoalexins but, significantly, it is about 300 times more potent than the lectin wheat-germ agglutinin, whose inhibitory effect on the growth of *Trichoderma* had previously led to the suggestion by Mirelman *et al*[5] that plant lectins (particularly those with binding specificity for *N,N′*-diacetylchitobiose and *N,N′,N″*-triacetylchitotriose) may act as natural antifungal agents

Schlumbaum *et al*[4] have re-evaluated

the antifungal activity of potential chitin-binding lectins Their results indicate that wheat-germ agglutinin, when prepared by the method of Mirelman *et al* [5], is contaminated with chitinase and has antifungal activity Four commercial preparations of wheat-germ agglutinin, which were not contaminated in this way, could not inhibit the growth of *T viride* Unfortunately, wheat chitinase is not recognized by antibodies raised against the bean enzyme, and it was not therefore possible in this case fully to exclude the involvement of other minor components (possibly even other lectins) in the antifungal activity of the preparation The demonstration of chitinase contamination in antifungal lectin preparations from tomato, potato and pokeweed, its absence from all lectins not showing antifungal activity, and the total inhibition of the antifungal activity of the tomato and potato lectins by anti-bean chitinase antibody strongly suggests that lectins do not inhibit fungal growth But they still have the potential to function in plant–microbe recognition associated with, for example, the binding of symbiotic bacteria to legume roots or the reception of polysaccharide elicitor molecules whose binding may trigger responses which include the synthesis of phytoalexins and chitinase

The studies of Schlumbaum *et al* [4] have relied upon the non-pathogenic *T viride* as the test organism, it is therefore now important to consider the effects of bean chitinase on true bean pathogens such as rust (*Uromyces phaseoli*) and anthracnose (*Colletotrichum lindemuthianum*) Induced resistance to the latter organism comprises, in addition to induction of chitinase, the accumulation of several isoflavonoid phytoalexins, hydroxyproline-rich glycoproteins and cell-wall-bound phenolic compounds[1], and the relative importance of these individual components to the overall outcome of the host–pathogen interaction is as yet unknown

A simple, direct test for the antifungal contribution of chitinase *in vivo* is not obvious, although the enzyme can be induced by ethylene, the relation between ethylene production and pathogen-induced defence responses is still unclear, and the use of inducers and inhibitors of ethylene production has nearly always yielded equivocal results[6]

Chitinase complementary DNAs and genes have recently been cloned from both plant (R Broglie, unpublished, C J Lamb, personal communication) and microbial[7] sources, and it may therefore soon be possible to assess the effects of heterologous chitinases on the disease resistance of transgenic plants It should be remembered, however, that chitinase alone may not have sufficient activity, Schlumbaum *et al* [4] refer in their paper to recent unpublished work suggesting that mixtures of chitinase and 1,3-β-D-glucanase (which are often co-induced in plants) may be more effective

The induction of chitinase in infected bean plants may be a direct response to the action of elicitor molecules A polysaccharide elicitor from the cell walls of *C lindemuthianum* can induce the enzyme in cultured bean cells[8] It has been known for some time, however, that chitosan, a partially depolymerized, de-acetylated form of chitin, acts as an elicitor of phytoalexin synthesis in peas and can also inhibit the growth of the pea pathogen *Fusarium solani*[9] The rapid accumulation of chitosan in infected pea pods would clearly require the action of constitutive chitinase and *N*-acetylglucosaminidase More recently, chitosan and chitin oligomers have themselves been implicated as elicitors of defence-related lignification in wheat[10] Future work will tell whether more than one chitinase gene exists in some plants, or whether the different functions of chitinase are exclusive to different host–pathogen interactions In the meantime, the potential importance of a naturally occurring, inducible antifungal agent arising as a single gene product should not be overlooked □

Richard A Dixon is in the Department of Biochemistry at Royal Holloway and Bedford New College, University of London, Egham, Surrey, TW20 0EX, UK

---

1 Dixon, R A *Biol Rev* **61**, 239–291 (1986)
2 Halverson, L J & Stacey, G *Microbiol Rev* **50**, 193–255 (1986)
3 Etzler, M E *A Rev Pl Physiol* **36**, 209–234 (1985)
4 Schlumbaum, A , Mauch, F , Vogeli, U & Boller, T *Nature* **324**, 365–367 (1986)
5 Mirelman, D , Galun, E , Sharon, N & Lotan, R *Nature* **256**, 414–416 (1975)
6 Chappell, J , Hahlbrock, K & Boller, T *Planta* **161**, 475–480 (1985)
7 Taylor, J D , Muller, G , Bond-Nutter, D & Jones, J D G *Hoppe Seyler's Z Biol Chem* **367**, (suppl) 148 (1986)
8 Dixon, R A *et al Phil Trans R Soc* (in the press)
9 Hadwiger, L A , Beckmann, J M & Adams, M J *Pl Physiol* **67**, 170–175 (1981)
10 Ride, J P *Hoppe Seyler's Z Biol Chem* **367**, (suppl), 75 (1986)

Retroviral immunization

Problems with AIDS vaccines

Peter Newmark

CURRENT hopes for a vaccine against AIDS (acquired immune deficiency syndrome) focus on the use of the whole or part of the envelope protein of the human immunodeficiency virus (HIV) as the protective antigen It is too soon to say whether or not this approach will succeed but the signs were not encouraging at a meeting held in Paris at the end of October* Participants were left doubting that envelope protein alone can induce the antibodies that would provide protection against the virus, suspecting that cell-mediated as well as antibody-mediated immunity will be needed, and worried about the availability of an adequate animal model and sufficient animals on which to test potential vaccines

At present, chimpanzees are the only animals considered to be suitable for testing AIDS vaccines They do not develop an immune deficiency as a result of exposure to HIV-1 but at least there is a persistent viral infection which should be prevented by any effective vaccine But first, any potential vaccine needs to be tested in chimpanzees for its ability to stimulate the immune system sufficiently for there to be a good prospect of protection — yet there are insufficient chimpanzees for all the testing that is and will be necessary The solution, according to Donald Francis (State of California Health and Welfare Agency) is for the initial tests to be carried out in humans His suggestion is that the immune system's response to any potential vaccine is first tested on human volunteers, who might particularly be found among health-care workers and monogamous homosexuals The vaccines inducing the strongest response would then be tested in chimpanzees for their ability to protect the animals against infection with the virus Success at that state would warrant vaccination trials in humans that fall into 'at-risk' groups

Meanwhile, at least one very small-scale trial of a potential vaccine in chimpanzees is nearing completion (Shiu-Lok Hu, Oncogen, Seattle) Two animals were immunized with a recombinant vaccinia virus carrying the HIV-1 envelope protein gene Not only did the animals raise antibodies against the protein but there was also evidence of cell-mediated immunity in the form of proliferating T cells and cytotoxic T cells (which released radiolabelled chromium from target cells infected with the recombinant vaccinia virus The animals have now been challenged with the AIDS virus and the outcome is awaited with interest

If this trial fails — and nobody was betting on a successful outcome — it will be particularly bad news in view of the observed immune response, which is the first to include the generation of cytotoxic T cells It seems unlikely that antibodies alone will provide an effective measure against viral infection and yet most potential vaccines other than the recombinant vaccinia virus induce antibodies but no clear cell-mediated response Even the antibodies they induce are of limited value in that they are usually unable to neutralize strains of the virus that are not closely related to the immunizing strain (Robin Weiss, Institute of Cancer Research, London, Dani Bolognesi, Duke University Medical Center) To mimic the more broadly neutralizing antibodies present in AIDS patients, it will probably be necessary to use a multivalent vaccine, incorporating the envelope proteins of several different strains of HIV-1 and perhaps also viral proteins that are not exposed on the surface of infected cells and yet might be important for the generation of a cell-mediated response (as in the case of influenza virus — see, for example, Ronald Germain's *News and Views* article in *Nature* 322, 687, 1986)

---

* *Retroviruses of Human AIDS and Related Animal Diseases*, Marnes-La Coquette, France 29–30 October 1986

There will be additional problems in designing a vaccine that provides protection not just against HIV-1 but also against HIV-2 (otherwise known as LAV-2 or HTLV-4) This virus, which is only rather distantly related to HIV-1, is the predominant HIV in western Africa, where it has been found both in association with AIDS (Luc Montagnier, Pasteur Institute) and in many symptom-free people (Myron Essex, Harvard School of Public Health) Essex reported 25 cases where both HIV-1 and HIV-2 were present as well as evidence that cells first infected with HIV-2 can still be infected with HIV-1, albeit with somewhat reduced cytotoxic effects Complementary data came from Patricia Fultz (Centers for Disease Control, Atlanta), who reported that a single chimpanzee infected with HIV-1 was still susceptible to HIV-2

Because HIV-2 seems to be closely related to simian immunodeficiency viruses (SIVs), there is some hope that it will be much easier to have an effective animal model for vaccines against AIDS produced by HIV-2 than by HIV-1 Ideally, HIV-2 would be found to produce AIDS in rhesus monkeys or some other species whose availability is not a problem But first results are not promising no virus could be isolated from rhesus monkeys following their infection with HIV-2 (Fultz, Ronald Desrosiers, New England Regional Primate Research Center) In the case of two baboons infected with HIV-2, virus was recovered for the first 3–4 weeks but not in the next 8 weeks (Desrosiers)

An alternative approach is to use a SIV as a model for HIV Indeed, Murray Gardner (University of California, Davis) reported that a killed vaccine was successful in protecting rhesus monkeys against simian AIDS But the virus that causes the disease is a type D retrovirus and, as such, is unrelated to the other human and simian viruses that cause AIDS, all of which are type C retroviruses Thus the general relevance of this success is in doubt (see Jerome Groopman's *News and Views* article in *Nature* 323, 489, 1986)

Indeed, an initial equivalent trial with a type C SIV was not successful (Desrosiers) In that trial two groups of four rhesus monkeys were vaccinated with an inactivated form of the type C SIV isolated from rhesus monkeys For one group the virus was given with threonyl muramyl dipeptide (a promising new adjuvant which was used in Gardner's successful vaccine), for the other group the vaccine consisted of viral proteins in the form of an iscom (immune stimulatory complexes in which the proteins are attached to a matrix) Eighteen weeks after vaccination the monkeys were challenged with the SIV Despite the fact that both vaccines induced a reasonable antibody response (although it was not determined whether neutralizing antibodies were present), when the monkeys were challenged with the rhesus SIV, the virus became established in the animals with a viraemia that has persisted for fifteen weeks, which is the duration of the experiment so far

For now, anybody who states that an AIDS vaccine will be available in 3 or 5 or 10 years time is simply hazarding a guess But reports of at least partly successful vaccines against two type C retroviruses — a recombinant vaccinia vaccine against Friend murine leukaemia virus (Bernard Moss, National Institutes of Health) and an iscom vaccine against feline leukaemia virus (Bror Morein, Swedish University of Agricultural Sciences, Uppsala) — suggest that the problem can be solved ☐

Peter Newmark is deputy editor of Nature

DNA structure

A simple solution to the stability of the double helix?

Maxim Frank-Kamenetskii

Why are different forms of DNA stable under different ambient conditions? This question becomes more and more important as the structural polymorphism of DNA (the left-handed Z form as well as the right-handed A and B forms) and its biological significance is more generally recognized. But is there a simple answer? Wolfram Saenger *et al* elsewhere in this issue (*Nature* 324, 385–388; 1986) claim that there is.

Their answer is the principle of hydration economy. In the B form of DNA the phosphate groups are far apart and are independently hydrated. In the A and Z forms the phosphate groups are brought closer together and are bridged by water molecules. Thus, in contrast to the B form, the A and Z forms are economically hydrated. As a result the latter two forms are stable under a water deficit, whereas the former is stable under a water excess.

This simple answer is based on much work in recent years in the field of DNA X-ray crystallography. Until recently X-ray diffraction studies of DNA were restricted to DNA fibres. In those studies water molecules were completely obscured. The structure of the DNA hydration shell emerged only after X-ray crystallography was applied to DNA fragments by A. Rich, R. Dickerson, O. Kennard and their co-workers in the 1980s. The first concept based on these data was the water spine, which is situated in the minor groove of B-DNA and stabilizes its canonical structure (Grew, H.R. and Dickerson, R.E. *J molec Biol* 151, 535–556; 1981). Now Saenger *et al* introduce the concept of hydration economy.

This novel concept is based on the fact that in the A and Z forms the free oxygen atoms of phosphate groups are, respectively, 5.3 and 5.1 Å apart and are bridged by water molecules. In B-DNA the distance between the oxygen atoms is 6.6 Å, too large for a water molecule to bridge. Hence the major conclusion of Saenger *et al* is that the A and Z forms should become more advantageous compared with the B form under a water deficit.

This concept readily explains the long-known fact that DNA dehydration (through drying or substituting alcohol for water) induces the B–A transition In poly(GC) poly(GC), dehydration induces the B–Z transition Saenger *et al* hypothesize that the mechanism of the famous salt-induced B–Z transition in poly(GC) poly(GC) is the same, as the salt pumps water molecules out of the polynucleotide.

Not everybody will share the latter viewpoint. Many believe that the high-salt B–Z transition is governed by electrostatic interactions of phosphate groups with each other and with the salt ions. In such a polyelectrolyte model, water does not play a structural role and only dampens the interactions between charges Doubtless the principle of hydration economy is just another factor in a very complex balance of interactions that governs the stability of double helical conformations. But its importance can no longer be ignored. ☐

Maxim Frank-Kamenetskii is in the Institute of Molecular Genetics at the Academy of Sciences of USSR, Moscow 123182, USSR

Celestial mechanics

Orbital evolution of comets

Julia Heisler

A FLAW in our understanding of the orbital evolution of comets is that the number of short-period comets — those with orbital periods less than 200 years, such as comet Halley — is much greater than theory predicts The discrepancy is enormous, the observed number is two orders of magnitude larger than expected[2] Relying partly on some newly recognized perturbations acting on comets, Bailey[3], on page 350 of this issue, presents a possible solution to this problem

In 1950 Oort proposed that the Solar System is surrounded by a spherical cloud containing 10^{11} comets, today it is believed that a typical 'Oort cloud' comet orbits the Sun at a distance of 25,000 astronomical units (AU) Until recently it was thought that the orbits of comets are mainly perturbed by random stars that pass within about one parsec of the Solar System In the past year, a second, more important perturbation source has been recognized The relatively smooth gravitational potential of the galactic disk in which the Solar System is embedded is responsible for a tidal effect upon comets that is proportional to the total density of material (stars, gas and dark matter) in the solar neighbourhood (This is analogous to the orbital precession of the Moon because of the pull of the Sun)

Both perturbations — from passing stars and galactic tide — gradually alter the orbital angular momentum carrying some comets into the inner Solar System and under the gravitational influence of the outer planets Those comets that pass within 5 AU of the Sun are observable from the Earth and are called 'new' comets, as they are probably making their first passage through the planetary system A new comet is recognized by its extremely small orbital energy (in absolute value), after one passage through the inner Solar System the orbital energy is typically increased sixfold Oort cloud comets that pass within 10 AU of the Sun are either ejected from the Solar System entirely or captured into short-period orbits However, even using both perturbations, the number of short-period comets derived from this capture process is still a small fraction of the observed number

A solution to this problem is that the short-period comets come from a source other than the Oort cloud, from a second, closer comet reservoir Current ideas[4] are that a comet belt extends outwards from the planetary system and departs from the ecliptic to blend in with the spherical outer Oort cloud at about 10,000 AU Comets at distances between a hundred and a few thousand AU are virtually unperturbed by stars, galactic tides and planets The comets from the inner region drift into the planetary system with perihelia well outside the detectable limit of 5 AU, and thus no significant flux of observable new comets is produced (In the case of the outer Oort cloud, the perturbations are stronger and some comets can be directly perturbed into orbits that are observable from the Earth) As a result of the lack of observable new comets from the inner Oort cloud, the spatial (or energy) distribution is very uncertain, and *ad hoc* assumptions are unavoidable However, Bailey[3] turns the short-period comet problem around, and demonstrates that the number of short-period comets constrains the choices, allowing one to give a crude determination of the spatial distribution

Bailey assumes a spherically symmetric cloud with a power-law energy distribution of unknown index He demonstrates that up to about 30,000 AU the galactic tide perturbs far more comets into the planetary region than stellar encounters do, and hence one can neglect stellar perturbations when calculating the injection rate into the planetary system This conclusion has been checked independently by numerical experiments[5] Bailey computes the fraction that will evolve to short-period orbits and multiplies that fraction by the average lifetime of observed comets to derive the expected total number of short-period comets To explain both the numbers of new and short-period comets, the energy distribution must be fairly steep For a comet lifetime of 3,000 years and an inner Oort cloud radius of 3,000 AU, 97 per cent of the comet population must be inside 10,000 AU This conclusion implies that most short-period comets originate in orbits near the cloud's inner boundary, whereas new comets originate at distances greater than 10,000 AU The short-period comets come from the inner Oort cloud and the new comets from the outer Oort cloud Furthermore, using entirely analytical methods and without invoking any other perturbations aside from the galactic tide, Bailey is able to derive the correct number of observed short-period comets

A massive inner cloud has also been invoked to explain biological extinction events on the Earth as a result of comet showers The related suggestion that large impact craters and extinctions over the past 250 million years occur periodically every 26–30 million years has received much attention, but there are doubts that the claimed periodicities in the data are genuine[6,7] It is also unlikely that comet showers are responsible for either periodic or random events on such short timescales Periodic comet shower theories include the proposals of a solar companion (Nemesis)[8,9], a tenth planet (Planet X)[10] and catastrophic encounters with giant molecular clouds during galactic plane crossings[11] The difficulty with the galactic plane crossing suggestion is that although the Sun does pass through the galactic midplane every 30 million years, the amplitude of its oscillation is small, about 70 parsecs Giant molecular clouds have a similar scale height, and therefore encounters can occur at all times during the Sun's oscillation, giving no statistically significant periodicity[13] In fact, despite the Sun's motion, the density of material near the Sun is fairly constant, implying that the comet flux induced by the galactic tide is also fairly uniform Finally, even random comet showers caused by close stellar encounters are unlikely to explain the crater record because an encounter capable of creating a shower is expected only once every 100 million years[5] Encounters with giant molecular clouds occur even less frequently, about once every 500 million years Hence, the crater data may reflect only a few comet showers This conclusion is supported by identifications of impact material at eight large craters up to 200 million years old, only two of the craters could have been formed by comet impacts, the rest were probably made by asteroids[14]

If Bailey's bold extrapolation of the total comet population is correct, the short-period comet problem is solved, it is satisfying that this problem is soluble using perturbation effects we already understand However, our theories about the inner region of the Oort cloud are still speculative, and we must proceed carefully, because unlike the situation for the classical Oort cloud, there are no observable new comets that directly sample the proposed inner comet cloud □

1 Joss, P C *Astr Astrophys* **25**, 271–273 (1973)
2 Fernández, J A & Jockers, K *Rep Prog Phys* **46**, 665–772 (1983)
3 Bailey, M E *Nature* **324**, 350 – 352 (1986)
4 Hills, J G , *Astr J* **86**, 1730 – 1740 (1981)
5 Heisler, J , Tremaine, S & Alcock, C *The Frequency and Intensity of Comet Showers from the Oort Cloud* preprint (1986)
6 Hoffman, A *Nature* **315**, 659–662 (1985)
7 Tremaine, S *The Galaxy and the Solar System*, (eds Bahcall, J & Smoluchowski, R) (University of Arizona Press, Tucson 1986)
8 Davis, M , Hut, P & Muller, R A *Nature* **308**, 715–717 (1984)
9 Whitmire, D P & Jackson, A A Jr *Nature* **308**, 713–715 (1984)
10 Whitmire, D P & Matese, J J *Nature* **313**, 36–38 (1985)
11 Rampino, M R & Stothers, R B *Nature* **308**, 709–712 (1984)
12 Weinberg, M D , Shapiro, S L & Wasserman, I *The Dynamical Fate of Wide Binaries in the Solar Neighborhood* preprint (1985)
13 Thaddeus, P & Chanan, G A *Nature* **314**, 73–75 (1985)
14 Weissman, P *Nature* **314**, 517–518 (1985)

Julia Heisler is at the Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

Role of HTLV III/LAV envelope protein

SIR—Sodrosky *et al*[1] recently reported that the envelope protein of HTLV III/LAV, the virus that causes AIDS (acquired immune deficiency syndrome), might cause the lysis not only of infected T-helper cells (T-4), but also of uninfected T-4 cells. This would explain the total depletion of T-4 cells, even though only some of them are infected with the virus. I would like to draw attention to another possible role for the protein in the pathophysiology of AIDS.

Developmental stages of AIDS are characterized by a selective major histocompatibility complex II (MHC II) restricted immunodeficiency which shows a striking resemblance to class I graft-versus-host immunodeficiency in mice[2,3]. This immunodeficiency is immunologically mediated by parental T-4 cell recognition of class II MHC alloantigens on the host's B-lymphocytes and results in a polyclonal proliferation of these B-cells and the formation of autoantibodies against a variety of polymeric antigens and also against envelope proteins of murine leukaemia virus bound on kidney cells[4]. These autoimmune processes lead to different pathophysiological symptoms, such as the selective immune deficit and autoimmune glomerulonephritis[4].

In the case of AIDS these antibodies might be especially focused on the envelope protein of HTLV III/LAV on T-4 cells. The antibodies against the envelope protein might initiate an autoimmune destruction of T-4 cells so that a vicious circle between T-4-cell immunodeficiency and T-4 directed autoimmunity can develop. This autoimmune destruction of T-helper cells might synergistically amplify the direct T-helper cell impairment by LAV/HTLV III and its envelope protein.

It is noteworthy that other autoimmune phenomena in patients with AIDS also show a striking resemblance to those in class II graft-versus-host disease in mice[4] and some human connective tissue diseases[5].

I suspect that not only classical cellular and antibody-mediated cytotoxicity[6] but also polyclonally activated transformed B cells might play a major role in the pathophysiology of AIDS. Strategies which inhibit this polyclonal B-cell proliferation might therefore delay T-4 deletion in early AIDS.

UWE BICKER

Boehringer Mannheim GmbH,
Sandhofer Strasse 116,
6800 Mannheim 31, FRG

1 Sodrosky J *et al* *Nature* **322**, 470-474 (1986)
2 Shearer, G M & Moser, M *Immun Today* **7**, 34-36 (1986)
3 Moser, M , Iwasaki, T & Shearer, G M *Immun Rev* **88**, 135-151 (1985)
4 Gleichmann, E *Immun Today* **5**, 324-332 (1984)
5 Lane, H C *et al* *New Engl J Med* **309**, 453-458 (1983)
6 Klatzmann, D & Montagnier, L *Nature* **319**, 10 11 (1986)

What is an extended probability?

SIR—While grateful for your open-minded discussion[1] of my review of extended probabilities[2], I should correct one minor point. I introduced the term 'extended probability' not (only) to avoid the charge of sensation-mongering but because probabilities exceeding unity are closely connected to, and in fact a consequence of, negative probabilities. Hence, the former have to be taken into account — if the latter "exist".

I confess my desire that somebody should give meaning to the term extended probability, but considering that Leonardo of Pisa in the thirteenth century was the first to interpret negative solutions of algebraic equations as debts, and even Descartes in the seventeenth century called them "false solutions" (the common term at that time), I do not believe this will happen in the near future. I have to disagree with the definition of negative probability given by D E Parry[3]. If, for instance, extended probabilities are used to obtain a formal resolution of the Einstein–Podolsky–Rosen paradox[4] they have to obey the following rule: an event leading with negative probability to a certain result has to extinguish not only the observation of a (hopefully) preceding exhibition of this result, but has also to extinguish the observation of this event at all, that is it has to reduce the number of accumulated events. (This is why a negative probability cannot be interpreted as a positive probability for observing some extinguishing counter-result, but only as a positive probability for the reduction of the number of preceding events.) And, of course, extended probabilities can only then creep into the picture when we deal with sets of indistinguishable results, as in the domain of quantum mechanics.

Unfortunately, no known physical mechanism allows for this kind of interaction or, therefore, provides a local and realistic resolution of the Einstein–Podolsky–Rosen paradox in terms of extended probabilities. Thus, the paradox will presumably continue to bother 'realistic' physicists, perhaps until this species becomes extinct.

W MUCKENHEIM

Sudring 10,
D-3400 Gottingen, FRG

1 Maddox, J *Nature* **320**, 481 (1986)
2 Muckenheim, W *Phys Rep* **166**, 337 (1986)
3 Parry, D E *Nature* **321**, 644 (1986)
4 *Lett Nuovo Chim* **35**, 300 (1982)

Migration and hominid bipedalism

SIR—Discoveries in palaeoanthropology have taken the story of hominid evolution back to a time just subsequent to the evolution of bipedalism[1]. The outstanding evolutionary question now is: what was the selection pressure that produced bipedalism? The current explanation[2] suggests that bipedalism evolved when the gathering of plants for food became necessary in a drier savannah habitat, and arms were needed to carry food back to a home site. Although this idea is feasible is it not compelling. We suggest a new idea in which bipedalism developed for long distance migration to scavenge from migrating ungulate populations, the only population that could provide sufficient food by scavenging. Bipedalism was a necessary adaptation to exploit this food supply.

In Serengeti, Tanzania, several ungulate species perform long-distance annual migration[3]. Migration provides an abundant and constant food supply, and hence these populations are large compared to sedentary populations of the same species. Mammal predators and scavengers do not follow the migrations because their young are slow growing and cannot travel with adults. Consequently breeding predators are sedentary. If a scavenger could migrate it would have access to an abundant and constant supply of carcasses (at least 1 carcass per 20 km^2 per day)[3], an order of magnitude greater than in non-migratory systems. In fact, humans can find enough to eat by following the wildebeest migration on foot[4]. Almost 70 per cent of carcasses are the result of undernutrition[3,5], so hominids would not be displacing predators from their kills.

Two vulture species follow the migrating ungulates, and they are considerably more numerous than sedentary species[5]. Vultures, however, have the disadvantage that they cannot easily break through the skin of intact carcasses. Therefore, there is in Africa an unfilled niche for a mammalian scavenger that can follow migrating ungulates, but such a mammal would need to carry its young.

We suggest that the selection pressure for the evolution of bipedalism in early hominids was access to this new rich and constant food supply. Morphs that had developed the ability to walk long distances would increase rapidly in number and displace the less numerous sedentary quadrupedal types dependent on plant gathering.

We envisage the protohominid quadruped as a plant gatherer and occasional scavenger, much like the baboon. For this type to follow migrating ungulates two essential adaptations are required simultaneously. First, members of the type must carry their young efficiently. This requires arms for carrying, and an upright stance. Chimpanzee females carry their young for at least the first year, so do humans, and thus it is reasonable to assume that early hominids also did so. Prolonged upright stance for efficient carrying would require adaptations of the hip joint, such as those in *Australopithecus afarensis* onwards[6] Second, they must travel long distances efficiently For this, the primitive foot with opposable big toe must change to the modern foot as a propulsive lever, and this change must coincide with changes in the hip

The alternative hypothesis is that bipedal hominids were plant gatherers in a savannah home range[2,6] Plant gathering required hands for carrying food, gradually the hominids' diet changed to include scavenged meat, using tools This idea is less compelling for two reasons First, modern baboons and chimpanzees are plant gatherers, and they can live in hot dry savannahs[7] Chimpanzees do not travel far (usually less than 4 km a day) Hence plant food gathering in dry environments need not lead to either bipedalism or long distance travel In contrast, efficient long distance travel (10–20 km a day) is necessary to follow the migrations If arms are used to carry young, bipedalism is the only efficient means[8] Second, a plant gatherer within a fixed range has few opportunities for finding ungulate carcasses, mortality of nonmigratory ungulates in Serengeti produces one carcass per month in 10 km^2 Thus selection to use tools would also be very small

The migration hypothesis suggests that tool use was an adaptation to speed up the butchering of carcasses and avoid competition with other stronger mammal predators Indeed, tools were used for scavenging 2 million years ago in Serengeti[9] The opportunities for hominid migrations to evolve were numerous, for savannah Africa was dominated by migration ecosystems[5]

The two hypotheses are testable and can be distinguished by the timing of changes in the foot For migration, hip and foot changes are both linked and essential Plant gathering, however, does not predict simultaneous changes in the foot because long-distance movement is not essential Unique hominid foot characters have been identified[10], for example the length of the foot is less than 25 per cent of the leg In future fossil finds, particularly for the period when bipedalism was evolving, it is such characters that will distinguish between the two hypotheses

We thank the New York Zoological Society and the Canadian NSERC

A R E SINCLAIR

Department of Zoology,
University of British Columbia,
Vancouver,
British Columbia V6T 1W5, Canada

MARY D LEAKEY

Box 30239, Nairobi, Kenya

M NORTON-GRIFFITHS

Ecosystems Ltd,
Box 21791, Nairobi, Kenya

1 Leakey, M D & Hay, R L *Nature* **278**, 317 323 (1979)
2 Isaac, G L in *Early Hominids of Africa*, 219 254 (Duckworth, London, 1978)
3 Sinclair, A R E & Norton Griffiths, M *Serengeti Dynamics of an Ecosystem* (University of Chicago Press, 1979)
4 Schaller G B & Lowther G R *Southwest, J Anthrop* **25**, 307 341 (1969)
5 Houston, D C in *Serengeti Dynamics of an Ecosystem*, 263 286 (University of Chicago Press 1979)
6 Lovejoy C O *Science* **211**, 341-350 (1981)
7 McGrew, W C *et al J Hum Evol* **10**, 1–18 (1981)
8 Fedak, M A & Seeherman, H J *Nature* **282**, 713-716 (1979)
9 Shipman, P *Am Anthrop* **88**, 27-43 (1986)
10 Lessertisseur, J & Jouffroy, F K in *Primate Functional Morphology and Evolution*, 327 340 (Mouton, The Hague, 1973)

---

^{131}I in milk and rain after Chernobyl

SIR—Radioactivity in rain water attributed to the Chernobyl plant was first detected in Japan on 3 May 1986, a week after the accident[1], and included ^{131}I We have monitored ^{131}I both in milk sold in markets and in fresh milk from dairy cows We found that the concentration of ^{131}I in milk in Nagoya, Japan at a distance of 8,000 km from Chernobyl was 4–5 times higher than that in rain water, and that both concentrations decreased with approximately the same effective half-life

Rain water was collected by a 0 305 m^2 collector between 3 May and 24 June The rain water with added carrier was concentrated by evaporation Gamma-ray spectroscopy showed the principal radionuclides to be ^{131}I, ^{103}Ru, ^{134}Cs and ^{137}Cs The ^{131}I concentration ranged from 980 pCi l^{-1} observed on 3 May to 0 019 pCi l^{-1} on 24 June, decreasing with an effective half-life of 5 2 ± 0 3 days (Fig 1)

The Ministry of Science and Technology announced that, in milk between 4 May and 12 May, ^{131}I activity ranged from 57 pCi l^{-1} to 400 pCi l^{-1} fresh milk and from 20 pCi l^{-1} to 160 pCi l^{-1} for commercial milk The northern jetstream which may have carried radioactive materials to Japan, passed over Nagoya, in central Japan We accordingly began a survey of the ^{131}I concentration in milk from five local dairies Because the ^{131}I concentration varied from dairy to dairy, no doubt due to the variety of feed furnished to cows, we decided to monitor fresh milk from individual cows mainly fed on fresh grass

Fresh milk sampled from 20 dairy cows at the university farm affiliated to the School of Agriculture of Nagoya University, was monitored between 19 May and 14 July Sixteen cows were of Holstein breed, and four were first-generation crosses between the Japanese breed and the Aberdeen Angus The cattle were each fed on fresh Italian ryegrass (34–35 kg per day) and concentrates (5–6 kg per day) The total milk volume from 20 cows amounts to 300–330 litres a day One litre weekly samples were placed in plastic bottles and mounted on a Ge(Li) semiconductor detector for gamma-ray spectroscopy ^{131}I was the only radionuclide identified in almost all of the samples, but a few gave extremely low ^{137}Cs signals For the last two samples, ^{131}I was chemically extracted and measured with a well-type Ge detector The ^{131}I concentration decreased with a half-life of 5 0 ± 0 2 days (Fig 1), similar to the half-life in dairy milk reported in Britain[2]

The ^{131}I concentration in milk was about 4–5 times greater than that in rainwater between 19 May and 14 July, a rainy season in Japan Judging from the ^{131}I concentration in rainwater and from that in milk announced by the ministry, it is reasonable to assume that the concentration in fresh milk reached a maximum on 4 May and declined with a half-life of 5 days

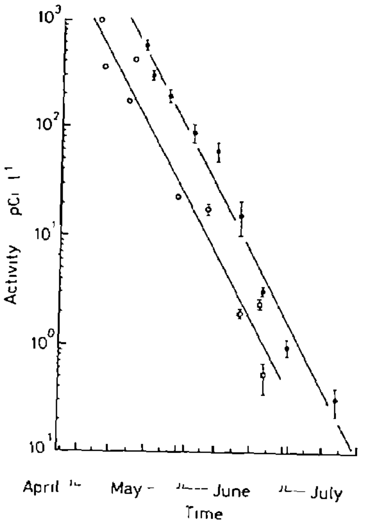


Fig 1 The ^{131}I concentration in dairy cow milk (black dots) and rain water (open circles) in Nagoya from 26 April, the date of the Chernobyl accident, to 20 July 1986

KUNIHIDE NISHIZAWA
KENZO TAKATA
NOBUYOSHI HAMADA
YOSHIMUNE OGATA
SADAO KOJIMA

Radioisotope Centre,

OKITSUGU YAMASHITA
MITSUAKI OHSHIMA
YOSHIMASA KAYAMA

School of Agriculture,
Nagoya University,
Chikusa, Nagoya 464, Japan

1 Aoyama, M *et al Nature* **321**, 819–820 (1986)
2 Hill, C R *et al Nature* **321**, 655–656 (1986)

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character They need not arise out of anything published in *Nature*, but those that do should not be highly technical comments on Articles or Letters (where the Matters Arising section remains appropriate) ☐

In Stock Now!

nature

INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

CALENDAR FOR 1987

ORDER YOUR OWN COPY NOW WHILE STOCKS LAST

featuring .
12 of our best cover photographs
in full colour

(Actual size
420 × 300mm)

Please send me **Nature's 1987 calendar:**

☐ One copy at £4 00 (UK) £5 00 (REST OF WORLD)
☐ Three or more copies at £3 50 (UK) £4 50 (REST OF WORLD) each No of copies ____________
☐ I enclose a cheque (made payable to **Nature**) for ____________*

**Payment may be made in your own currency at the prevailing rate of exchange*
Prices include postage and packing

Signature ____________
Name ____________
Address ____________

Send orders to Circulation Manager, **Macmillan Journals Ltd**, FREEPOST (UK only), Brunel Road, Houndmills, Basingstoke, Hants RG21 2XS UK

JOURNALS FROM MACMILLAN

☐ *ANTI-CANCER DRUG DESIGN*

NEW JOURNAL FIRST ISSUE PUBLISHED OCTOBER 1985

(Quarterly)

Editors S Neidle S T Crooke

The development of new clinically-useful anti-cancer agents is increasingly dependent on rational scientific principles drawn from a wide range of disciplines This new journal aims to promote such approaches, providing a common meeting ground for medicinal chemists, pharmacologists, biochemists, biophysicists and oncologists involved in both fundamental and clinically orientated aspects

This journal is published with the support of The Cancer Research Campaign

NEW JOURNAL FIRST ISSUE MAY 1986

☐ *BONE MARROW TRANSPLANTATION*

(Quarterly)

Editor J M Goldman Co-Editor R P Gale

A new journal which will publish clinical results in man and experimental results with animals, allografting and autografting, HLA and other methods of selecting donors, graft-versus-host disease and all types of methods to prevent and treat it, and much more

☐ *BRITISH JOURNAL OF CANCER* *(Monthly)*

The Journal of the Cancer Research Campaign

Chairman D G Harnden Editor M Moore

One of the leading cancer journals in the world publishing articles on all the major scientific disciplines

Publishes papers from researchers throughout the world and the proceedings of meetings of the British Association for Cancer Research and other symposia

☐ *HUMAN TOXICOLOGY*

(Bi-monthly)

Secretary to the Editorial Board· P Turner

Covers all aspects of human toxicology embracing both animal research with direct relevance to man and studies with humans

SAMPLE COPY ORDER FORM

Name ____________
Position ____________
Address ____________
____

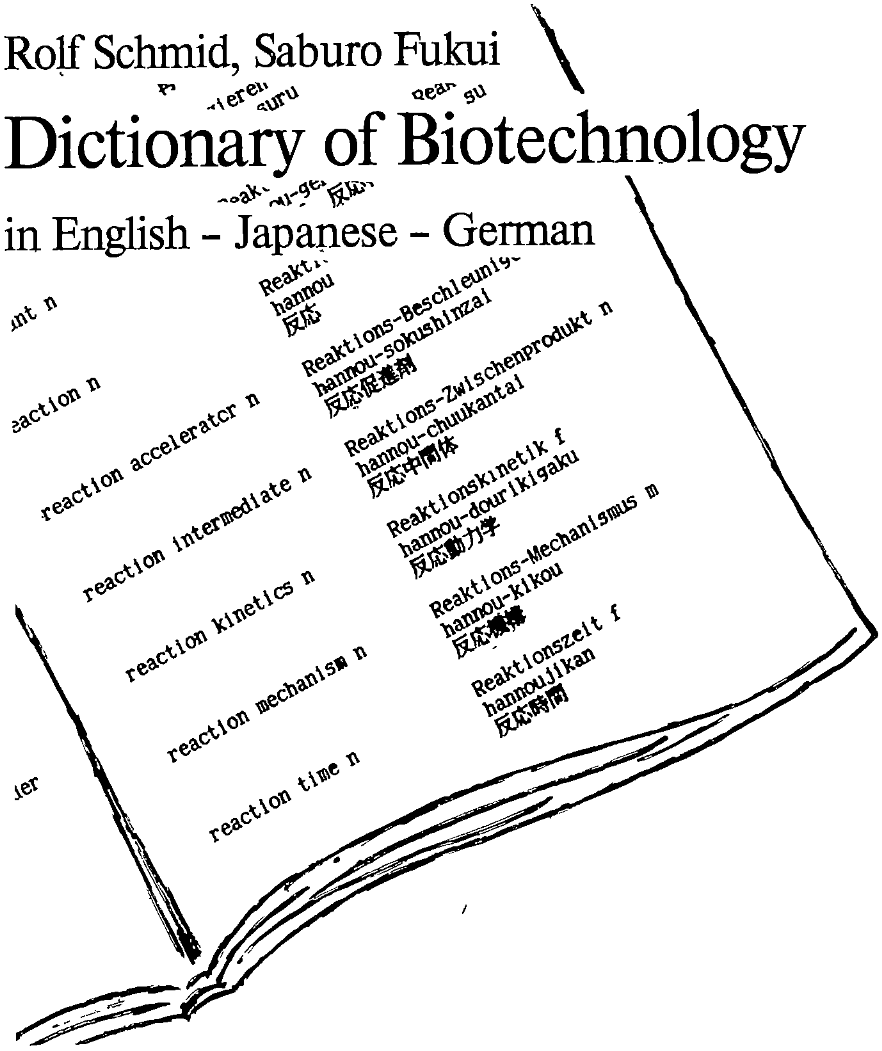
Please tick the sample copies required and return to Alison Baverstock, Macmillan Press, Houndmills, Basingstoke, Hants, RG21 2XS, UK

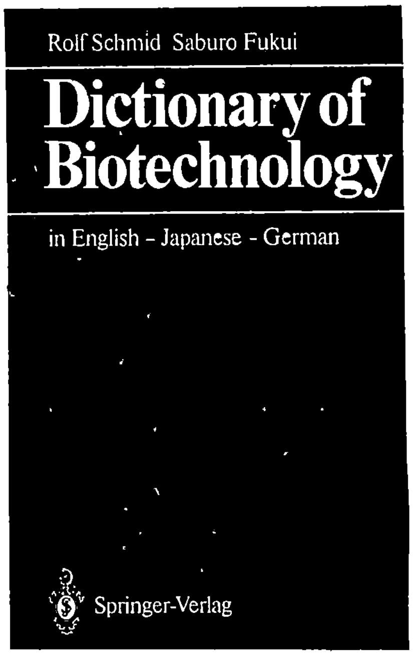
Macmillan Press
Scientific & Medical

Rolf Schmid, Saburo Fukui

Dictionary of Biotechnology

in English - Japanese - German




Rolf Schmid Saburo Fukui

Dictionary of Biotechnology

in English - Japanese - German

Springer-Verlag

With a foreword by
Bundesminister für
Forschung und Technologie
Dr. Heinz Riesenhuber

1986 XV, 1324 Seiten
Hard cover DM 298,-
ISBN 3-540-15566-X

Contents: Technical Terms English-German-Japanese - List of Synonyms. - Technical Terms Japanese-English-German - List of Synonyms - Technical Terms German-English. - List of Synonyms - Appendix.

This "Dictionary of Biotechnology" is the outcome of several years' work with English, German and Japanese texts in this field. It contains nearly 6 000 terms and is divided into three parts 1. Alphabetically listed technical terms in English accompanied by translations into German and Japanese, 2. Alphabetically listed technical terms in Japanese accompanied by translations into English and German, 3. Alphabetically listed technical terms in German accompanied by translations into English
Each of the main parts is followed by a list of synonyms in the word entry language An appendix defines over 300 abbreviations using all three languages The English terms are classified according to n(noun), vb (verb), and adj (adjective). The spelling of German scientific terms is tending to adapt to Anglo-American patterns The German terms are followed by an indication of gender (m, masculine, f, feminine; n, neuter, pl, plural). For the sake of easier legibility, the Hepburn method has been selected for transcribing the Japanese terms into the Roman alphabet ("roma-ji"). The division of lengthy transcribed terms into roots conforms with a didactic approach

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo

Heidelberger Platz 3 D 1000 Berlin 33
175 Fifth Ave, New York, NY 10010 USA
28, Lurke Street Bedford MK40 3HU, England
26, rue des Carmes F 75005 Paris
37 3, Hongo 3-chome Bunkyo-ku Tokyo 113, Japan

7994/4/1

Springer

Reader Service No 6

The patient's dilemma

D.J. Weatherall

Unnatural Selection? Coming to Terms with the New Genetics. By Edward Yoxen.
William Heinemann: 1986. Pp. 208. £12.95.

GEORGE Bernard Shaw once observed that all professions are conspiracies against the laity, a view that has permeated much of the bad press to which doctors have been subjected in recent years. What has gone wrong? Paradoxically, the successes of the medical sciences in the post-war period, which saw the development of antibiotics and the eradication of many childhood infections by vaccination programmes, may be partly to blame. These remarkable achievements raised the expectation that the control of cancer, heart disease, arthritis and other major killers would soon follow. When it became apparent that this was not to be, and that these conditions could only be managed, often unsatisfactorily, by increasingly sophisticated patch-up procedures involving the use of expensive drugs, a bewildering array of machinery and the replacement of almost every organ except the brain, a sense of disillusionment with scientific medicine set in. Hospitals came to be perceived as terrifying places to visit, let alone in which to be ill, and it was widely believed that the pastoral needs of patients had become lost in a maze of high technology. Alternative forms of therapy were sought, and "holistic" medicine, a term which simply means good doctoring, was rediscovered.

It is hardly surprising, therefore, that when terms such as *in vitro* fertilization and genetic engineering appeared on the scene further public suspicion was generated; the technocrats were at it again, now breeding monsters in test tubes and tinkering with people's genes. The situation was exacerbated by sensational television programmes which featured parents wandering around "gene supermarkets" stocking their baskets with traits that they hoped to see expressed in their children. Memories of Nazi Germany and the eugenics movement were revived, committees of inquiry set up, and hastily constructed and ill-thought-out bills put before the British Parliament in an attempt to prevent any form of research involving human embryos.

Considering the complexities of molecular and cell biology, and the speed with which these fields have advanced over the past few years, it is not surprising that the clinical applications of these new techniques have been greeted with apprehension; the medical profession itself is only just starting to get to grips with their potential. Clearly, a long period of public debate and reflection is required before we shall be in a position to assess how far we wish to go in interfering with our reproductive activities and manipulating our genomes. This new book by Edward Yoxen should be a valuable catalyst for pointing the discussions in a sensible direction.

Like Jane Austen and J.D. Salinger, Yoxen believes in the analeptic effect of a racy opening paragraph: "There is a great deal of sex in this book, quite enough to shock the faint at heart". After an introductory section (which rapidly disillusions even the not so faint at heart), successive chapters deal with artificial insemination, *in vitro* fertilization, surrogacy, sex predetermination, antenatal diagnosis, new forms of genetic intervention and gene therapy. In each of these controversial areas the author faces up to the ethical and pastoral problems that are posed and gives a balanced discussion of the benefits and risks involved. By and large he subscribes to the value of these developments, while at the same time showing great sensitivity to the views of those who find many of them distasteful. It is to be hoped that his sane account of such an emotionally-toned topic will be read by teachers, politicians, church leaders and, in particular, members of the many pressure groups that are attempting to prevent further research without evaluating the evidence.

Yoxen makes little attempt to explain the scientific basis of the advances in reproductive physiology and genetics that have given us these powerful new tools, and, in most cases, simply states what is or is likely to be possible and the moral dilemmas that may arise. Presumably he feels that this is as far as he can go at the moment, given the level of biological sophistication in society. Many will agree. In discussing the problem of scientific communication, Francis Crick has described how he had to explain to a lady visitor to the Salk Institute that worms are not spontaneously generated in apples. And, as Yoxen points out, it is not all that long since those of the British aristocracy who wanted male heirs submitted to hemicastration in the belief that sperm from the right testicle produce sons — *noblesse oblige* indeed.

Such tales may not be the best guide to the level of scientific knowledge in the world at large, but there is no doubt that if Yoxen's goal is to be reached — that of an individual who can say "I fully understand the step that I am taking, I have chosen it freely, having sought advice and considered the options available, and I recognize my responsibilities to myself and to others in what I do here" — we shall have to do some fundamental rethinking about education. To no country does this apply more than Britain, in which children are encouraged to specialize at the age of 14, and in many cases to ignore science completely thereafter, and which, with a few exceptions, is very badly served by its media in scientific matters. Thus far, prejudice and ignorance have characterized many of the views on the topics discussed in this book. Anyone who doubts that there is a problem should read *Hansard's* account of the parliamentary debate on embryo research.

There is another question that is touched on several times in Yoxen's book, but which is never fully developed. What is it that motivates scientists in this field? Why is it, for example, that in the United States there may well be nearly as many research workers trying to replace the defective gene in a rare genetic form of immune deficiency as there are patients with the disorder? And why has at least one worker jumped the gun in a premature attempt at gene therapy? Are patients becoming stepping stones on the way to scientific "firsts" and fame and fortune? As part of the dialogue that is essential if public confidence in modern medicine is to be restored, biologists and medical scientists will have to provide better insights into how science works. They should also ask themselves why it appears that so much talent is being directed into research that gives rise to increasing public concern when, at the same time, there seems to be much less interest in many of the more immediate medical problems of society. ☐

D.J. Weatherall is Nuffield Professor of Clinical Medicine at the University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK.

Taking the soft tomato

THE business of reviewing scientific books is one apart from the reviewing of works of literature. But anyone whose book has been roughly handled in a review can find comfort in *Rotten Reviews: A Literary Companion*, which contains such comments as this: "Here all the faults of *Jane Eyre* (by Charlotte Brontë) are magnified a thousand fold, and the only consolation which we have in reflecting upon it is that it will never be widely read". Thus one James Lorimer in a review of Emily Brontë's *Wuthering Heights* (1847).

Rotten Reviews, edited by Bill Henderson, is a 93-page, small-format volume which contains extracts from 175 scathing notices of books that have proved to be of enduring appeal. Publisher is Pushcart Press, and distributor W.W. Norton, New York. Price is $12.50.

Beastly creations

Martin Kemp

The Rhinoceros from Dürer to Stubbs, 1515–1799. By T. H. Clarke. *Sotheby's Publications/Philip Wilson, Russell Chambers, Covent Garden, London WC2E 8AA, UK: 1986. Pp. 219. £29.50.*

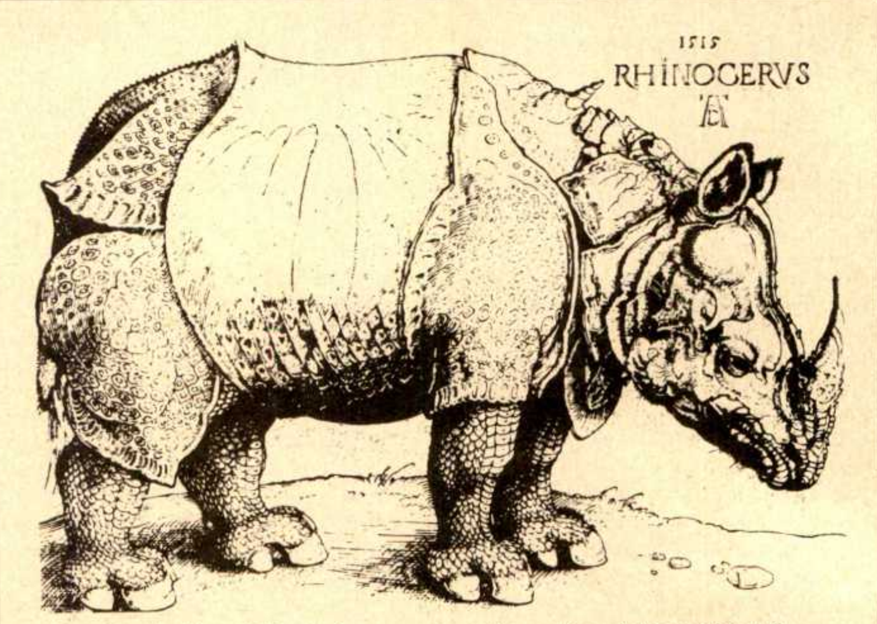
RHINOCEROTICA rules. We might expect an obsessional study of images of a single exotic animal over three centuries to be rich with informative detail. We would not necessarily expect it to be fun. T. H. Clarke's book is both.

The author does not attempt to rival the scholarly detail of L. C. Rookmaaker's *Bibliography of the Rhinoceros* (A.A. Balkema, Rotterdam; 1983). Rather, he has provided a lively compendium of artistic images, ranging from soberly naturalistic records to elaborate fantasies in all kinds of media — painting, sculpture, drawing, engraving, embroideries, tapestry, pottery, porcelain, glass, furniture, and arms and armour. The text is written with a nice wit that only occasionally becomes tiresome.

The rhinoceros is unusually well suited for such a study. It was a rare enough sight for European observers to be the subject of wonder and even fantasy, yet there are enough documented instances of rhinos on European tours to provide a series of identifiable animals from which the great majority of the images sprang. Clarke discusses eight peripatetic visitors between 1515 and the 1790s. In addition to the host of craftsmen who took advantage of the rhinomania occasioned by exhibition of the animals by showmen, a number of accomplished artists were involved. To mention only Albrecht Dürer, Hans Burgkmair, Pietro Longhi, Jean-Baptiste Oudry and George Stubbs will give some idea of the quality and the geographical range.

The involvement of considerable artistic talents in naturalistic illustration has obvious advantages. What is not so readily recognized is that their sheer skill brings dangers as well as benefits. The kind of artistry developed during the Renaissance revolution in representation — and so vividly manifested in Dürer's art — could be used to endow an image of fantasy with a sense of "naturalistic" credibility to match that of a representation taken from life. A skilful artist can make a dragon look as credible as a rhinoceros for someone who has seen neither.

This kind of potency in "fantastic naturalism" explains the immense hold of Dürer's engraving, which, as Clarke indicates, was based on secondary visual and written sources. Dürer's transformation of the leathery rhinoceros into a fearsome piece of late-mediaeval siege machinery resulted in an image of such power that it remained truer to the emotional response evoked by the animal than the more naturalistic versions. Clarke provides an abundant body of visual evidence to show that naturalistic illustration is a double-edged weapon, though he does not grapple head-on with this fascinating question.



Dürer's rhino, 1515 — "a fearsome piece of late mediaeval siege machinery".

From *The Rhinoceros from Dürer to Stubbs*

The greatest strength of the book is the remarkable range of media, particularly from the applied arts, that is illustrated and discussed with authority. Its two most obvious weaknesses relate to the history of natural history on one hand, and iconographic art history on the other. Historians of natural history will be disappointed that the great succession of zoological encyclopaedias is so little exploited, while art historians might reasonably have expected a fuller discussion of the emblematic and symbolic meaning, including the sexual connotations of the "horn". These are not separate matters; for the encyclopaedist, nature was full of meaning, just as meaning for the emblematicist was regularly borne in a natural vehicle. Clarke's approach largely excludes such issues, but it does permit a wide measure of visual delight. ☐

Martin Kemp is Professor in the Department of Art History, University of St Andrews, Fife KY16 9AL, UK.

Nervous activity

D. Regan

Evoked Potentials. Edited by Roger Q. Cracco and Ivan Bodis-Wollner. *Alan R. Liss/Wiley: 1986. Pp. 551. $96, £70.*

OUR knowledge of the cellular and neuropharmacological organization of the brains of experimental animals has advanced enormously over the past decade. How this new knowledge can be used to deepen our understanding of human vision, hearing and motor control, and how it can help to prevent and alleviate visual, hearing and spinal cord disorders, is an immediately challenging problem. It is unfortunate that many of the important advances in neuroscience have been achieved through techniques that, for ethical reasons, can seldom be used in man. A major role of evoked potential recording is to bridge the gap between the human brain and the results from cellular-level experiments on animals.

The term "evoked potential" (EP) covers the electrical responses of the entire peripheral and central nervous system, including peripheral nerve action potentials, spinal cord responses to stimulation of peripheral nerves, the responses of the auditory brainstem, midbrain and cerebral cortex to sounds, the responses of somatosensory brainstem and cerebral cortex to tactile stimulation of the skin, and the responses of the retina and visual cortex to light. The term also covers the electrical activities of the brain associated with the evaluation of the meaning of a stimulus, such as a spoken word, rather than with the physical nature of the stimulus. In this book the "evoked potential" takes on an even wider meaning: Williamson and Kaufman describe how the magnetic responses, 10^{-7} weaker than the Earth's field, that are produced from neurones in the depths of the brain can be clearly picked up by a quantum interference device placed near the head, and the location of the neurones defined to within a few millimetres!

Evoked potentials can be recorded

from metal electrodes held onto the scalp or over the spine without discomfort or risk. Although they are very small (often less than one microvolt), a domestic microcomputer is quite adequate to extract the EP from the electrical "noise" produced by muscle and by the rest of the brain. Evoked potentials thus offer access to the electrical activity of the human brain's visual, auditory and tactile areas while the brain's owner is describing and measuring what he or she sees, hears and feels. In this way the technique can help us to understand the neural activities that underlie normal sensory experience. By the same token, EP recording is a tool for investigating the neural abnormalities associated with disorders of sight, hearing and motor control in human adults, infants and neonates.

This book contains a selective collection of 49 admirably concise mini-reviews on a wide variety of topics, each written by an internationally recognized authority. Illustrating how EPs can be used functionally to "dissect" a sensory pathway, Bodis-Wollner's opening chapter describes the parallel visual subsystems for pattern, flicker, motion and so on. Arguably, the most promising clinical role of the EP technique is in studying disease mechanisms in paediatric patients. After pointing out that the early identification of hearing loss can avoid the disastrous consequences of speech delay, Galambos, and Stockard and Stockard, go on to review the recent advances in identifying hearing defects in newborns by EP recording. Sokol, and Levi and Manny, describe how visual EP recording can aid the management of infants with amblyopic visual loss and help to prevent lasting visual impairment. The possibility of using spinal EPs to assess spinal cord trauma incurred during birth is reviewed by Cracco and Cracco. The applications of EPs in neurology and ophthalmology are extensively discussed, including a useful analysis of visual evoked potentials in multiple sclerosis by Blumhardt and three papers on Parkinson's disease and neurotransmitter abnormalities. The sections on surgical monitoring and on psychiatric disorders in adults and children are instructive and thorough. The book contains an excellent group of papers on the neural sources of sensory and cognitive EPs.

The editors are to be congratulated on putting together such a coherent survey of basic and clinical human electrophysiology. All 49 chapters are carefully and concisely written, and are remarkably free from errors. A uniformly high standard is maintained throughout, and several chapters are quite outstanding. This is the best and most up-to-date book on the topic. ☐

D. Regan is a Professor in the Departments of Ophthalmology and Medicine, Dalhousie University, Gerard Hall, 5303 Morris Street, Halifax B3J 1B6, Canada.

Developing policies

Ronald Coleman

The Biotechnological Challenge. Edited by S. Jacobsson, A. Jamison and H. Rothman. *Cambridge University Press: 1986. Pp.181. £22.50, $34.50.*

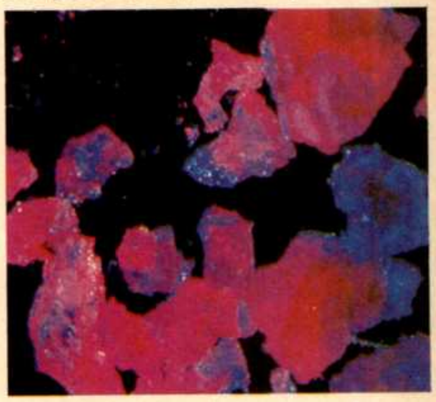
MANY nations, both developed and developing, have attempted to define strategies for the application of biotechnology to improve their overall national economic performance over the next 10 to 20 years. Any such strategy must take account of the science base of the country concerned, the availability of trained manpower and the state of other, complementary technologies, as well as the overall national industrial performance and the ability to deliver goods to the marketplace. This book offers advice on the potential of biotechnology for developing countries.

The editors have brought together experts in their respective fields to outline the current status of the technology, and also present some case-studies relevant to developing countries. For policy-makers, the techniques underpinning modern biotechnology — genetic engineering, and enzymes for biotransformations and fermentation — are well presented in adequate detail. But the choice of case-studies could have been better, and it is unfortunate that none of them relates to health care.

One of the studies deals with the production of high-fructose corn syrup. This is the largest scale industrial use of modern biotechnology, and it has been successful in the developed world, particularly the United States, because the syrup can compete successfully, both in performance and price, with a traditional agricultural product (in this case sugar). But in a developing country sugar production must surely remain the preferred route for the supply of sweeteners: with corn syrup, the possibility of competing successfully for export markets is remote, and it makes no sense to import the necessary technology for a new sweetener in order to achieve a marginal improvement for home markets.

The account of the Brazilian experience in the conversion of sugar cane into alcohol, as a replacement for petroleum, is much more illuminating. Simple economics suggest that the process is less efficient than selling sugar on the world market and using the income to import petroleum. Given the current constraints on Brazil, however, the strategy seems to be both appropriate and successful. The natural advantages of an equatorial country — abundant sunshine, rain and cheap labour — have been used effectively to incorporate a biotechnological process into a large national programme to develop an alternative energy source. Part of the Brazilian experiment is the forward integration of the programme so that the alcohol derived from sugar is used in specially designed engines produced internally, further reducing the outflow of foreign currency. There is little prospect of much competition from manufacturers in the developed world, and the possibility of export opportunities to other countries with a capability of producing alcohol but with a less developed industrial base.

The discussions of biotechnology in food production, particularly aimed at an increase in protein for developing countries, are dated and do not give adequate recognition to the current surpluses of carbohydrates everywhere in the world except Africa. The use of simple technologies, such as the immediate drying of harvested materials to prevent spoilage by fungi, is likely to give a much better return than investment in advanced biotechnological techniques.



Colour display — the picture (of interferon crystals) is reproduced from Biotechnology: Strategies for Life, *by Elizabeth Antébi and David Fishlock, a large-format, highly-illustrated popular book newly published by MIT Press. Price is $39.95, £29.95.*

It is an especial pity that the book does not include a concluding chapter by the editors, with advice on the prerequisites of a successful process or industry based on the biological sciences. The technology itself is only a small part of the overall picture, and at this stage is probably best imported through licence agreements or joint ventures with leading companies. In the longer term, the commercialization of biotechnology will require substantial investment in trained manpower, and the development of a broad-based research and development programme, and industrial production facilities. Meanwhile, the best hope for the developing countries must be in forming suitable partnerships with other nations. *The Biotechnological Challenge* contains this message, but in only muted tones. ☐

Ronald Coleman is the Government Chemist in the Department of Trade and Industry, Cornwall House, Waterloo Road, London SE1 8XY, UK.

Distillation to the basics

D.H. Watson

Fundamental Virology. Edited by Bernard N. Fields and David M. Knipe. *Raven: 1986. Pp.784. $54.50.*

JENNER may seem to have been anachronistically prescient in writing of smallpox *virus*. Of course, he only used "virus" in the then-conventional sense of something unpleasant (literally venom or poison). The more restrictive use in the present century was promoted by the observations of Iwanowski and Beijerinck that the causal agent of tobacco mosaic passes through a filter capable of impeding bacteria. Similar properties were soon demonstrated for some organisms infecting animals, followed by recognition of bacterial viruses which, despite their later discovery, were to be the main focus of the early years of virology. A notable landmark was the introduction of physical methods and, in particular, of a rigorous quantitative approach, through the recruitment to the discipline of some distinguished physical scientists (whose scientific emigration has sometimes been attributed to their post-war disenchantment with aspects of modern physics).

Real advance in animal virology only came when animal cells could be cultivated with an efficiency approaching that of their bacterial counterparts. Most importantly, this led to Dulbecco's development in 1982 of reproducible plaque assays for animal viruses. Succeeding years saw accelerating progress in animal virology, which both benefited from and contributed to the growing power of molecular biology.

More recently, it has become clear that the molecular approach to the study of viruses causing animal and, in particular, human disease can lead to rational approaches to chemotherapy, new kinds of vaccines and useful insights into mechanisms of pathogenesis. As a result, modern textbooks on "medical" virology now refer to molecular aspects of the subject, as well as the more traditional clinical aspects of virus infections. The ultimate in this genre must be the encyclopaedic *Virology*, edited by Bernard Fields *et al.*, and published by Raven last year.

Fundamental Virology is a derivative ("exprint" perhaps?) of that earlier volume. It consists, in half the page extent of its parent, of the more basic, rather than the clinical or applied chapters.

The preface concludes with admirable confidence that the text will be useful for courses in general and molecular virology. I agree, but feel the editors may be a little too sanguine in anticipating that it will be used by senior undergraduates, although it will inevitably be the standard reference book at this level. It should certainly be used and studied at the postgraduate level, and by teachers and researchers.

It is perhaps a little harsh to cavil at the coverage of what is a large volume, even in the magnum, rather than jeroboam, version. Nevertheless, the title is a little misleading. Inevitably, from its provenance, the book deals with the fundamental virology of human viruses, whereas any course on molecular virology must by necessity address bacterial, plant and other animal viruses.

This apart, the volume covers its chosen field comprehensively and all of the chapters have been written by experts on their topics. No other book has such extensive coverage which is so up-to-date, with references up to 1984. □

D.H. Watson is a Professor in the Department of Microbiology, University of Leeds, Leeds LS2 9JT, UK.

Back to the future

John Treherne

Trillion Year Spree: The History of Science Fiction. By Brian W. Aldiss with David Wingrove. *Gollancz/Atheneum: 1986. Pp.511. Hbk £15, $19.95; pbk £9.95.*

BRIAN Aldiss and David Wingrove cast a wide net in their 500-page history of science fiction, a revised edition of Aldiss's *Billion Year Spree* published in 1973. Inhabitants of their pantheon range from Gilgamesh and Disraeli to Mickey Mouse and Ray Bradbury. Only poor Fred Hoyle, one of the few scientific professionals among British SF writers, is missing.

It all started in the halcyon years of the Regency, with Victor Frankenstein and his 19-year-old creator, Mary Shelley. Before Mary, there was only what the authors call "ur-scientific fiction" (that is, before the genre originated), with such promising "ur-SF" types as Dante, Voltaire, Daniel Defoe, Will Shakespeare, Erasmus Darwin and the writers of *Genesis*. After Mary, come the real SF-pioneers — Edgar Allen Poe and Bram Stoker, sustained by gothic nightmares of nasty things in dark airless chambers, headless corpses and Count Dracula; followed by Jules Verne, H.G. Wells and Conan Doyle — with giant space rockets, fearful weapons, squidgy creatures in long-legged machines on Cobham Common and the Earth ploughing through toxic interstellar gas — and Edgar Rice Burrows (Tarzan) and William Hope Hodgson, with swine-things that come up from the cellar while the Sun goes dull in the sky.

Meanwhile, "post-ur" auxiliaries were obligingly scribbling away in the wings: Samuel Butler and William Morris (with their Utopias); Thomas Hardy (whose astronomer-hero, Swithin St Cleeves, gazes at the night sky with Lady Constantine wondering if "horrid monsters lie up there?"); Robert Louis Stevenson (Dr Jekyll, *of course*); even the gentle, consumptive Richard Jefferies (imagining the worst in *After London*) and many, many more. All are lovingly recruited to the SF cause and accorded lengthy quotation.

Then on past Freud and the Nazi party; Čapek and Kafka, robots and *RUR*, the destruction of civilization; Aldous Huxley and the imbecility of industrial progress; Hollywood and C.S. Lewis; to the whole tottering panoply of late-twentieth-century SF with its intergalactic wars and time warps, bug-eyed monsters (still), supermen and genetic disasters, occasional romps on the wilder shores of psychology and pharmacology and, most surprising of all, the return to pure gothic fantasy with goblins and wizards, absolute evil, sandworm gods and beautiful princesses in Middle Earth, Dune and Gormenghast.

For most of us, I suspect, encounters with SF are limited to youthful passions for Flash Gordon, Dr Who or Captain Kirk and, in adult years, the occasional Asimov and Aldiss snatched in desperation at Heathrow or from a railway bookstall. The first requirement for the uncommitted SF reader — suspension of disbelief — can be difficult, especially for those of a pedestrian turn of mind. It is not only the convoluted plots and scientific improbability, there are all those silly names. Even Doris Lessing becomes a challenge in this scenario of one of her SF novels:

> Two mighty Galactic Empires, one — Canopus — spiritual, one — Sirius — technological, are interested in the degenerated planet Shikasta, once a paradise world called Rohanda. Agents from Canopus and Sirius walk its surface and witness stages in its development and decay...

But for Aldiss and Wingrove this is a serious business:

> Science fiction is the search for a definition of mankind and his status in the universe which will stand in our advanced but confused state of knowledge (science), and is characteristically cast in the Gothic or post-Gothic mode.

Undoubtedly they are right when the writers are Wells or Huxley or Kafka and, perhaps, to some extent even when they are L. Ron Hubbard or Arthur C. Clarke. Yet I suspect that most SF readers, like their Edwardian predecessors, are after a rattling good yarn in the realms of ultimate fantasy. And this SF clearly provides in great abundance, as is charted in this readable, scholarly and — despite much good humoured bravado — slightly apologetic book. □

John Treherne, a scientist and novelist, is at Downing College, Cambridge CB2 1DQ, UK. His most recent novel is Mangrove Chronicles *(Jonathan Cape, 1986).*

Andrómaco

a breakthrough
in
immunology

AM3* ORAL BRM

(BIOLOGICAL RESPONSE MODIFIER)

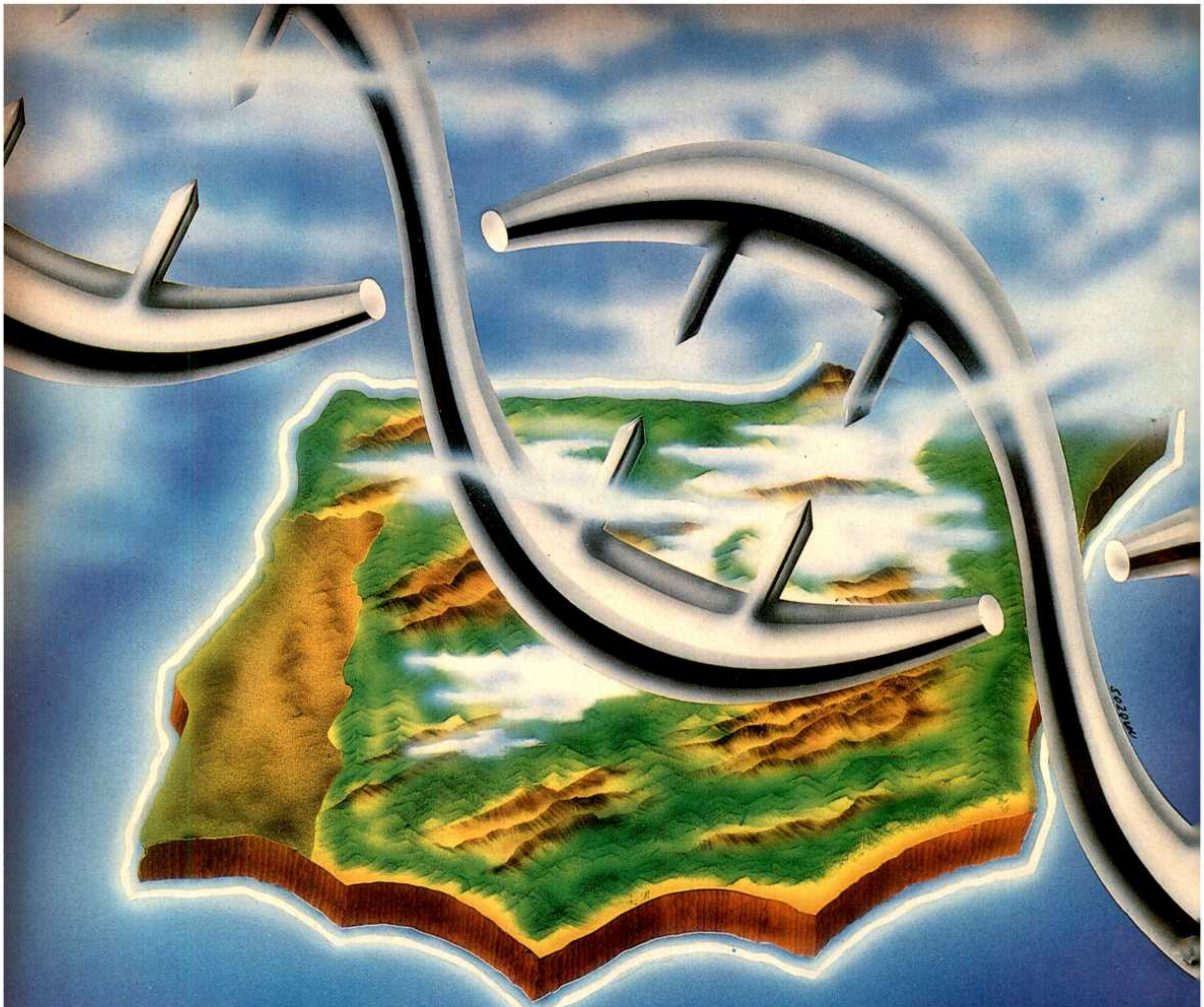
MAKES PHARMACOLOGICAL BONE MARROW ENHANCEMENT POSSIBLE

(*) polysaccharide - protein complex

For further information please contact

Adolfo BASILIO
Director of Licensing and new products
Laboratorios ANDROMACO
C/. Azcona, 31 - 28028 Madrid SPAIN
Phone 255 56 00 Telex 49 325 LACE

Reader Service No.3



The missing link

CULTEK is a Spanish Company founded in 1976, specialised in the distribution of instruments, reagents, chemicals and disposables for the biotechnologist. We represent a number of well-known manufacturers from around the world (i.e. Bio Rad, Flow, etc.) and sell and give service to over 2.000 customers throughout Spain, working in the following areas:

- *Molecular Biology*
- *Celular Biology*
- *Immunology*
- *Specialised diagnostics.*

Over the last decade, we have been the pioneers in Spain of these new technologies, which gives us a leading position in terms of experience and know-how in league with the world's best. We bridge the gap between manufacturers and users, overcoming financial, bureaucratic and linguistic barriers, and serve as technical consultants for both partners in the chain. In summary, we are the missing link.

If you plan to market or buy your biotechnology products in Spain, do not waste time and effort, come to an expert.

Cultek

The experts in Biotechnology

CULTEK, S. L. Caídos de la División Azul, 20. 28016-MADRID. Telephone (91) 457 08 90. Telex 46.658 CLTK E.

Reader Service No.177

CULTEK'S HOUSE AGENCY.

A renaissance in the making

An OECD survey in 1984 placed Spain and Portugal among nations that "give little or no priority to research and development". But they now seem aware of the need to build scientific resources.

IT is hard to square the gloomy judgements of the OECD examiners with the reality of a visit to Iberia. Whether Madrid or Lisbon, science administrators are full of enthusiasm for the future and confident that much can be achieved. It is perhaps after all only human nature: things really are getting better and better and although research investment lags far behind Europe's major nations there is none of the despondency that comes with living in the declining north.

Wider horizons came with the ending of the dictatorships that had governed Spain and Portugal. Although a decade has passed since the first heady days of freedom ("when we talked politics 24 hours a day because we had never been able to before" as one researcher put it) there is still a sense that new opportunities are opening up. This year it is entry into the EEC.

Major challenges come with membership. Both Spain and Portugal are predominantly agricultural nations and lack powerful multinational companies. And while both receive enormous numbers of tourists each year (Spain considerably more than its own population), neither wants to see its involvement with Europe end in economic colonization by the wealthier north.

Names, money and thanks

IBERIAN names, as this correspondent found out, are rather complicated. Having signed my full name, Alun Mark Anderson, in a small hostel in Bilbao I found myself being addressed as "Señor Mark" for the duration of my stay. I might, I am told, have equally well have been addressed as "Señor Mark Anderson" or even "Don Alun" but no one would have committed the foolishness of calling me "Señor Anderson", my presumed maternal surname.

If surnames become muddled in the articles that follow, or if I have given some people more surnames than others, I apologize in advance for my ignorance. Doubtless there will be other errors for this survey is based on a trip to Spain and Portugal that was all too short. Whatever errors there are, they will not be the fault of the people who gave me special help in planning my trip and whose knowledge of Spanish (or Portuguese) science is much deeper than my own.

In Madrid special thanks go to Dr Paul S. Dick. Science Officer at the British Council and his staff for their hospitality and for arranging many of my visits. In Barcelona Dr Pedro Puigdoménech and Professor Juan Subirana deserve equal thanks for looking after every aspect of my visit. And in Lisbon various anonymous members of the Ministry of Education deserve thanks for arranging my schedule.

NOTE: £1 = 198 Spanish pesetas = 229 Portuguese escudos. □

The increase in agricultural and industrial productivity that both nations now urgently need is inevitably entwined with scientific and technological development. The scale of the problems faced by the two nations is, however, quite different. Spain has four times the population and six times the land area of Portugal. More important, it has a gross domestic product (GDP) per head of £3,087 (1983 figures) that puts it within shouting distance of those, like Britain (GDP per head, £5,542), at the back of the pack of advanced industrial nations. Portugal, with a GDP per head of £1,380, is still at an intermediate stage of development with parts of its agricultural system a century behind the rest of Europe.

Both countries have a history of spending appallingly little on research and development. Even the most optimistic estimates suggest that Spain spends 0.6 per cent of its GDP on research and development and Portugal 0.8 per cent, in comparison with Britain's 2.4 per cent.

Both countries inherited from their dictatorial past institutional structures ill-equipped to deal with rapid change. Their research organizations were fragmented, with different sections of activity separated from one another, and research separated from teaching in the universities. Arbitrary and autocratic control was exerted from above.

The first wave of change was catalysed by young and impatient men and women who had been trained abroad. Even before the dictatorships vanished, many of the brightest students had begun to go to foreign laboratories to gain doctoral degrees. The returnees now provide the core of many of the research groups of international standing and provide the impetus for change.

Spain has moved forward at some speed. Structures have changed and attitudes sometimes seem to be struggling to catch up. In the universities and research institutes, the psychology of the Franco era was expressed by autocratic professors presiding over miniature empires, and in concern with international politics and the skillful currying of favour over the cultivation and reward of excellence. The old Spain is still visible in the laboratories but it is dying, laid low by increasingly open competition for positions and grants.

Overall coordination of the whole spectrum of research is now being installed with a new law and a new interministerial body. National priorities will be set and national plans made. A centralized form of administration thus returns, but the government assures doubters that Franco's ghost will not be found within it: researchers have never had greater control over their day-to-day affairs. The new body will, however, have to choose the area in which grants will be given out, while still bearing in mind that the results of much of the research must eventually find application in Spain's agriculture and industry.

That this can be a difficult task is suggested by the failure of a national programme for biomass development launched by the new commission's predecessor. What makes picking winning areas particularly hard is the lack of clear signals from industry of the direction in which it is going. Research spending by industry has always been very low. The new commission will thus have the double task of promoting research that can find application while convincing industry that it should be taken seriously. The fear is that it will be easier to follow international fashions (biotechnology, electronics and new materials are bound to be the new litany) than to find areas in which Spain has a unique chance of success.

The pace of reform in Portugal has been slower. A perfectly good organization to coordinate science and technology exists, but only gradually is it being given the money and powers to do its job.

Meanwhile different research organizations and different sectors of the government happily go their own ways without any real coordination from the top and a curious university research system keeps hundreds of small groups isolated from one another without any real incentive to achieve excellence.

Portugal is saved from its lack of coherent planning by its size. The research community is still small enough for almost everyone to know everyone else, which often means that bureaucratic barriers are not so formidable in the laboratories themselves. Wisely, Portugal is concentrating efforts to develop its backward inner regions through the establishment of new universities and polytechnics with an emphasis on agriculture. Here the scope for improvement is so great that even a little technology transfer is likely to work miracles. □

Spain: political change

A high-speed ride to democracy

In June of this year the Spanish people re-elected a socialist government with one of the largest majorities in Europe. A little less than ten years earlier that party was illegal, defined itself as "Marxist...rejecting any path of accommodation to capitalism" and had a revolutionary programme intended "to socialize production, distribution and exchange by the working class".

In power the party has taken a very different line; it won a referendum to keep Spain in the North Atlantic Treaty Organization (NATO), took Spain into the EEC and is now busy selling off the vast state holdings accumulated during the Franco era to private industry. But that a radical party transformed itself into something like a conventional European social democratic party is but one of the wholly unexpected twists of history that has brought Spain, in ten years, from a fascist dictatorship to a parliamentary monarchy.

The Congress of Deputies takes cover as Lt Colonel Antonio Tejero Molina and the Civil Guard take over.

The roots of the peaceful transition can be found in the late 1950s when, with Spain close to bankruptcy, Franco allowed the economy to be liberated. Spain was opened to foreign investment and foreigners and as the tourists flocked in, hundreds of thousands of Spaniards went to work in the then booming industries of northern Europe. The results were dramatic; between 1961 and 1973, the "years of development", Spain grew faster than any country except Japan. By the time the boom was over in 1974, Spain found itself with a society much of which was middle class and which had a vested interest in the continuing stability of the country. In 1975 Franco died and a series of new actors appeared on the political stage.

Six years before his death Franco appointed his own heir (a prerogative he had granted himself). His choice — Juan Carlos, the grandson of the last king of Spain, Alfonso XIII, who been brought up under the tutelage of General Franco — was not a surprise. But the King was not to prove a mere follower to Franco. After Franco's death, he astonished the nation by appointing as Prime Minister, Adolfo Suarez, who had spent his life working for Franco and seemed totally unsuited to dealing with a nation demanding change. But the King's choice proved correct; Suarez's knowledge of the dictatorship was to prove the very strength that enabled him to dismantle it. Change followed at a giddy pace.

Suarez was appointed in May 1976; by November he had persuaded the Cortes, stuffed as it was with old supporters of Franco, to agree to the establishment of a two-chamber parliament and universal suffrage. In December 1977 the change was approved by a referendum; the following spring political parties were legalized, including even the communist party. In June the first general election was held. Not surprisingly, Suarez's own moderate party was the winner — but not by an overall majority. To the nation's astonishment (and the disappointment of those who thought Franco's death would herald a revolution) Suarez then proceeded to negotiate a comprehensive series of agreements, the Moncloa Pacts, with all the opposition parties and then, when this was done, to draw up by consensus a new constitution defining Spain as a parliamentary monarchy.

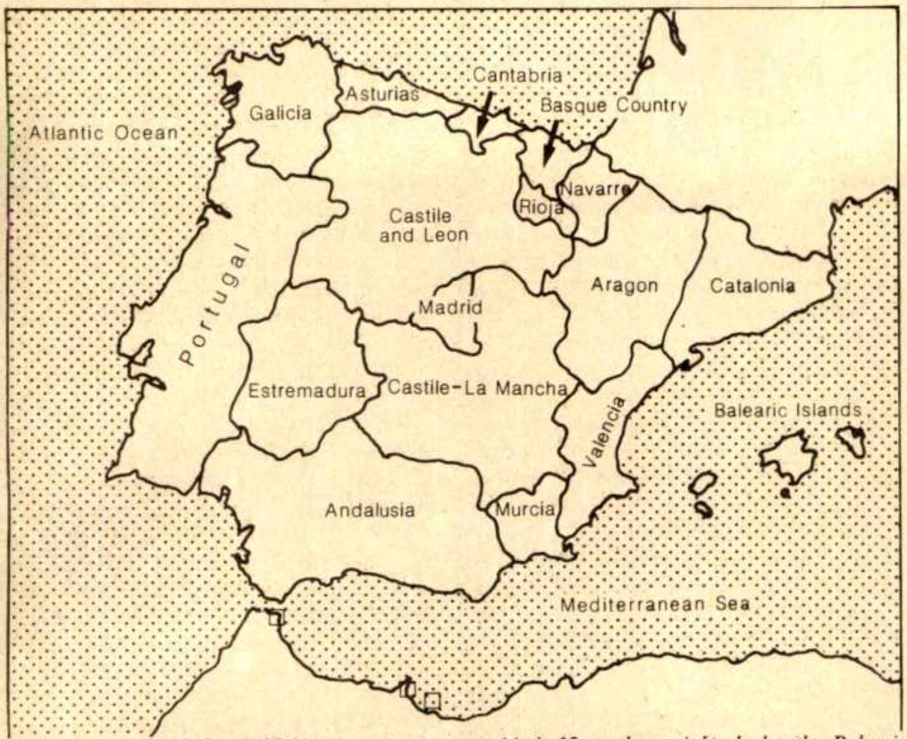
While it was that act in October 1978 that marked the official beginning of the new Spain, it was two and a half years later that democracy had its baptism of fire. On 23 February 1981 (a date which all Spaniards seem to know), with Suarez still in power but his own party increasingly divided, a group of senior army officers decided that democracy had gone too far. The Congress and virtually all of Spain's top politicians were seized by a group of Civil Guards. In Valencia tanks took to the streets. But the attempted military coup was to fail, largely through the prestige of the King. By remaining firmly opposed to the coup and persuading military leaders not to back it, King Juan Carlos was able to maintain control and peace was restored in twenty-four hours.

The following year the party that Suarez had led was in such disarray that he himself left to form a new one. That left the field open for the Socialist Party to win the general election with the 40-year-old Felipe Gonzalez as its leader. Just seven years after Franco's death Spain had an elected socialist government. Since then, Gonzalez has been skillfully piloting his party towards the political centre. After further success in the 1986 election he appointed the youngest cabinet in Europe, many in their early 40s and with a Minister of Industry and Energy who is just 35.

Alongside the establishment of democracy, other changes have taken place. Less well noticed abroad, but of great significance in Spain, has been the devolution of power to regional governments; Spain is half-way to becoming a federal state.

Devolution came in bits and pieces and was intended at first to provide for the demands of the Basque region — where violent demonstrations for independence had begun earlier — and Catalonia and Galicia; all regions with their own languages and culture. Spain now consists of 17 'autonomous regions'; each has its own ministers, president and administrative staff, but there similarities end. The Basques have gone furthest down the road to autonomy. They collect their own taxes and send on to Madrid their share (a little over 6 per cent) of the cost of 'central' services — defence, foreign affairs, aviation and the like — while keeping back taxes for local government, including health, housing, social welfare and agriculture. Catalonia, in contrast, is content to let the central government collect taxes and receive their share. The Basques also set up their own police force and both the Basque Country and Catalonia have their own television channels and bilingual road signs. These changes have helped to quieten separatists, but the problem remains; the main political wing of the Basque separatist movement has renounced violence, but its leftist military wing, whose own representatives are still able to win around 15 per cent of the Basque vote, has not.

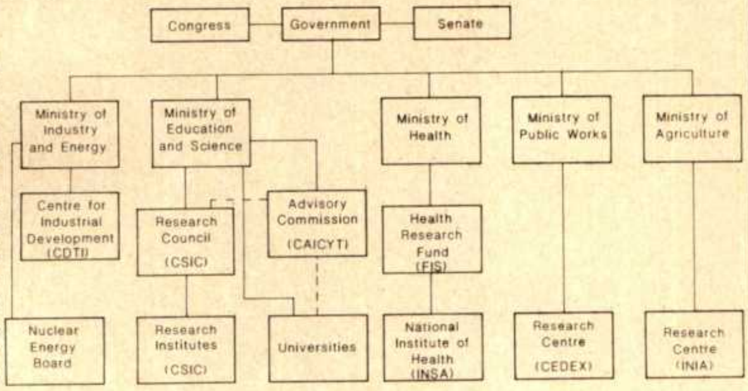
Science, technology and education remain a disputed area in which clear and universal transfer of authority has not been arrived at. The Basques and Catalans are responsible for the running of their universities and schools, but in about half the regions the universities remain firmly in the hands of the central government, for the time being at least. No region has been able to grasp full financial control of its science and technology budget, although both the Basques and Catalans have their own regional science administrations. The unwillingness of the central government to transfer these powers is still very much an issue of argument. ☐



Spain is now divided into 17 'autonomous communities'; 15 on the mainland plus the Balearic Islands in the Mediterranean and the Canary Islands in the Atlantic. With somewhat different status are the two tiny North African enclaves of Ceuta and Melilla located, aptly enough, just across the sea from Gibraltar. Morocco would like the two towns back.

Science bodies await a reshuffle

SPAIN's science is not run quite like this any more (see chart) for the new law for the "Promotion and general coordination of scientific research and technological development" is just beginning to make itself felt. Under the old system the Ministry of Education and Science looks after both the universities, which have not traditionally been active in research, and the Consejo Superior de Investigaciones Cientifica (CSIC), the science research council in whose various institutes research facilities are mainly concentrated. Grants for the research performed in both the universities and in the CSIC institutes are assessed by the Comisión Asesora de Investigacian Cientifica y Tecnica (CAICYT).

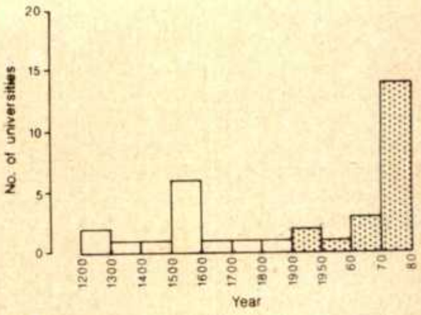

Congress
Government
Senate
Ministry of Industry and Energy
Ministry of Education and Science
Ministry of Health
Ministry of Public Works
Ministry of Agriculture
Centre for Industrial Development (CDTI)
Research Council (CSIC)
Advisory Commission (CAICYT)
Health Research Fund (FIS)
Nuclear Energy Board
Research Institutes (CSIC)
Universities
National Institute of Health (INSA)
Research Centre (CEDEX)
Research Centre (INIA)

The Ministry of Industry and Energy is responsible for nuclear power development and for industrial development through the Centro para el Desarrollo Tecnológica y Industrial (CDTI), an organization advancing money for research where the risk is too high to attract ordinary investors. Medical research is funded through a seperate granting agency, Fondo de Investigaciones Sanitarias (FIS). The Ministry of Public Works retains a large research centre, the Centro de Estudios y Experimentación (CEDEX).

Change came to the Ministry of Agriculture a few years ago when it lost direct control over most of its research centres in favour of the autonomous communities in which they were located. But it retains responsibility for overall coordination and for Madrid's Institute Nacional de Investigaciones Agrarias (INIA).

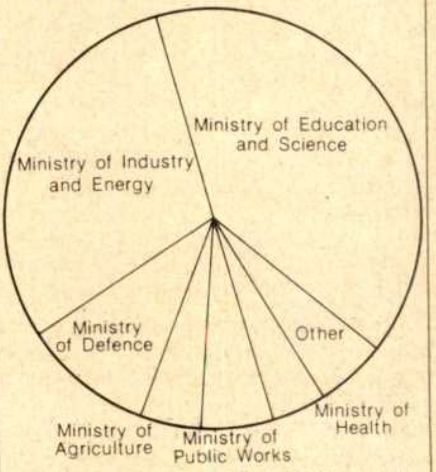
Under the new system the institutions survive but an attempt is made to improve coordination and planning through central control by the Interministerial Commission for Science and Technological Development which contains representatives from all the ministries. This will sit at the centre like an octopus with its tentacles reaching out to all the research agencies. Its responsibility will be to programme all their activities through the construction of national plans. Advice on the design of the plan, its objectives and its progress will fall on an Advisory Council (a transformation of CAICYT). Quite how the new system will function is not yet clear for the Interministerial Commission has only just met for the first time, a situation perhaps symbolized by new charts showing everything joined to everything else by dotted lines. □

Towards a second golden age



Why a second golden age? A glance at a chart showing when the universities were established shows that there has not been a boom like the past decade's since the sixteenth century. That was the time of the height of Spain's prosperity when riches were pouring in from the newly-colonized Americas (the Aztecs were conquered in 1519 and the Incas in 1532) and the King of Spain was also Holy Roman Emperor. Holland, Austria and much of Italy were under his control and Spain was the dominant power in Europe . Of the universities established then, Valencia, Santiago, Sevilla, Madrid (Complutense), Granada, and Zaragoza still survive. □

Education ministry is the big spender



THE Ministry of Education and Science is the big spender among the government agencies with responsibility for research, even bigger than the Ministry of Industry and Energy. Total funds expended by the public sector reached 119,000 million pesetas this year — 80 per cent of the total. The small contribution of the private sector contrasts strongly with industrial nations like West Germany, the United Kingdom and the United States, where more than half of all research is funded by industry. The discrepancy helps to explain why the percentage of Spain's gross national product spent on research is so low, at most 0.6 per cent, in comparison with the major industrial nations which score from 2.2 to 2.8 per cent. □

Law of science

More coordination from the top

"SCIENTIFIC research and technological development have traditionally taken place in Spain in a climate of apathy and lack of social stimuli, in the absence of the necessary apparatus to guarantee the efficient intervention of the authorities in the planning and coordination of the scant means available, with a lack of connection between the aims of research and the policies of the relevant sectors and, usually, between the research centres and producing sectors. It is hardly surprising, therefore, that Spanish contribution to scientific and technological progress has in general been small and inappropriate to the position Spain has occupied in other areas; and that, when this has not been the case, as in certain periods of the present century, the most worthy contributions have been the result of the isolated effort of particular individuals."

IT would be impossible not to applaud a government that is able to set down such a brutally frank appraisal of its own nation's shortcomings as that given above, taken from the introduction to the "Law for the Promotion and General Coordination of Scientific Research and Technological Development", passed by the Parliament earlier this year. Of course, the government finds it easier to make sweeping criticisms with the knowledge that the present defects can be blamed on the past dictatorial system. But even so the new law is a remarkable attempt to remedy the defects listed above, with the aim of ensuring that Spain can look forward to the day when it can speak on equal scientific terms with neighbouring countries.

The new law is not the beginning of the reform of scientific organizations in Spain. The ten years that have passed since the end of the Franco era "feel more like a century", as one researcher put it. The two largest concentrations of scientific talent, the institutes of the research council (CSIC), which have traditionally been the stronghold of research, and the universities, which have concentrated more on teaching, have both undergone considerable reform, as the following pages will make clear. What the new law really does is to reform the highest level of decision making and to bring it under the control of a new policy body answerable to parliament. That body is the Interministerial Commission for Science and Technology and it will be responsible for designing and overseeing the execution of the National Plan for Science and Technology.

The Interministerial Commission will contain representatives from all the ministries concerned with science and be chaired by the Minister of Education. It will have the power to decide the budgets, staff, and capital needed to carry out the national plan, to allocate funds to the various organizations involved in carrying it out and to coordinate the research activities of the different scientific bodies. Independent action by separate research organizations, each of them presenting its own budgets to the Ministry of Finance, will thus end. The plans of the commission will be debated by parliament and a yearly report on their progress made public.

It is not, however, the intention of the commission to impose rigorous control from the top over all aspects of scientific research. Within the national plan, there will be "National Programmes of Scientific Research and Technological Development" (in which information technology, biotechnology, and materials science are likely to feature strongly) designed by the commission. But the national plan will also incorporate "sectorial programmes" emanating from government departments and research bodies, and programmes from the autonomous communities. Overall control of the national supply of research personnel will be taken as the responsibility of the commission.

The status of a number of autonomous research bodies, including CSIC and the Nuclear Energy Board, will also be altered by the new law, principally so that they may more flexibly interchange staff.

Those who designed the new law feel proud of it; indeed they think the new structure could teach a thing or two to other countries, like the United Kingdom, which they see as very poor models for overall research planning. But it is much too early to say whether the Spanish system will succeed for, at the time of going to press, the commission will only have just had its first meeting. How quickly it will impose itself on the organizations it now commands remains to be seen.

Strangely, many in CSIC and the universities seem quite uninterested in the new law and doubt that it will bring about shifts in the direction of ongoing research. Perhaps more pertinent is the criticism

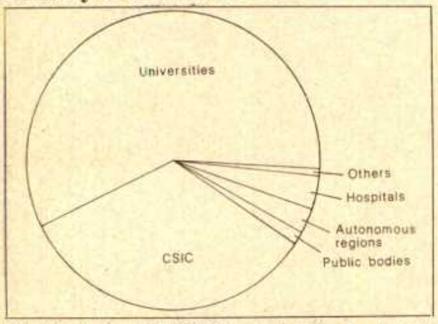
Sharing out the CAICYT cake

FOR the universities, CAICYT is the major source of research grants, although private foundations, industry and the Fondo de Investigaciones Sanitarias (FIS) of the Ministry of Health and Consumer Protection are important. The latter provides much support to medical faculties and hospitals. Those working in research council (CSIC) institutes also have grant applications assessed by CAICYT but their grants can be topped up from CSIC's own funds.

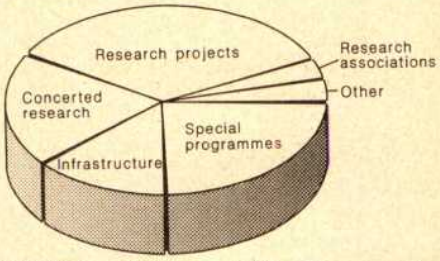
Supporting research projects uses up the largest slice of CAICYT's budget, but still accounts for less than half of total expenditure. Special programmes take a quarter of the funds. They come in two varieties, mobilizing programmes which aim at the general promotion of a whole field of science (current themes are high-energy physics and biotechnology), and national priority research and development programmes which also promote coordinated action across a whole field, but aim at a more concrete economic objective. Both involve all research organizations, particularly the institutes of CSIC. Four programmes have been launched in the latter category: biomass energy (which proved such a disaster that it was abandoned), aquaculture (which has stimulated the establishment of a new CSIC National Centre for Microelectronics with laboratories at the Autonomous University of Barcelona and the Polytechnic University of Madrid) and rail transportation.

Research in industry is promoted through concerted research programmes in which loans (up to 50 per cent of total cost) are given to private firms to cover their research costs. Part of the research may be totally funded if it is carried out at a government laboratory. If there are good prospects for development of a marketable product then help from CDTI (see p.324) is likely to come in the later stages.

An increasingly large slice of the budget goes to improve research infrastructure, working through the autonomous communities. Smaller sums go to aid industrial research associations, to encourage foreign scientists to take their sabbaticals in Spain and to support international cooperation, scientific publications and congresses. ☐



Distribution of CAICYT's project funds. A total of 3,826 million pesetas was given out in 1985.



CAICYT's total budget — 9,100 million pesetas in 1985, but 13,000 million next year — is widely spread.

K

~~C~~HROMATOGRAPHY

with K as

KONIK

NK-3000 HRGC

KNK-500 HPLC

DATA SYSTEM BASED
ON IBM® - PC (or compatibles)

INSTRUMENTS, S.A.

ecialists in chromatography and ancillary techniques

rnational Marketing Centre

ELONA: Crta. Cerdanyola, 65-67 - San Cugat del Vallés. 08190-BARCELONA (SPAIN)
Tel. (93) 674 32 50 - Telex 59.199 - Apdo. 136

es and Service in over 42 Countries
der Service No.176

10th
KONIK
ANNIVERSARY

PANREAC

- Reagents for Analysis, Purissimum and for Synthesis.
- Solvents for Spectroscopy UV–IR.
- Solvents for Liquid Chromatography.
- Analytical Volumetric Solutions.
- Buffer Solutions.
- Primary Chemical Class.
- Indicators.
- Reagents for Clinical Diagnostic.
- Special Reagents.
- Chemical-Pharmaceutical products according to pharmacopoeias.
- Auxiliary products for Laboratory.
- Pure Chemicals.

SV PA PS RE PRS CL

MONTPLET & ESTEBAN, S.A.

CHEMICAL MANUFACTURERS

Riera de Sant Cugat, 1
08110 MONTCADA I REIXAC
(Barcelona) Spain

Telephone: (93) 564 24 08*
Telex: 57485 PANR E

Center Regional Office:
Sierra de Guadarrama, 36
28850 S. FDO. DE HENARES
(Madrid) Spain
Telephone: (91) 656 64 12*
Telex: 46713 PANR E

aditio

Chemicals for the Food Industry according to specifications of the FOOD CHEMICALS CODEX.

Industry Authorized and Registered in the Ministerio de Sanidad y Seguridad Social, n.° 31.38/B

Reader Service No.4

Publications of
The ASTRONOMICAL
Society of Japan
—Bi-monthly—

Published by **Astronomical Soc. of Japan**

devoted to original papers in all the fields of astronomy & astrophysics including the following theoretical and observational categories.

- Astrometry
- Earth and planetary systems
- Celestial mechanics
- Galactic structure & stellar dynamics
- Galaxies and cosmology
- High-energy astrophysics
- Astronomical instruments & data processing
- Interstellar matter
- Physical processes & astrophysical plasmas
- Stellar atmospheres
- Stellar structure & evolution
- and the Sun

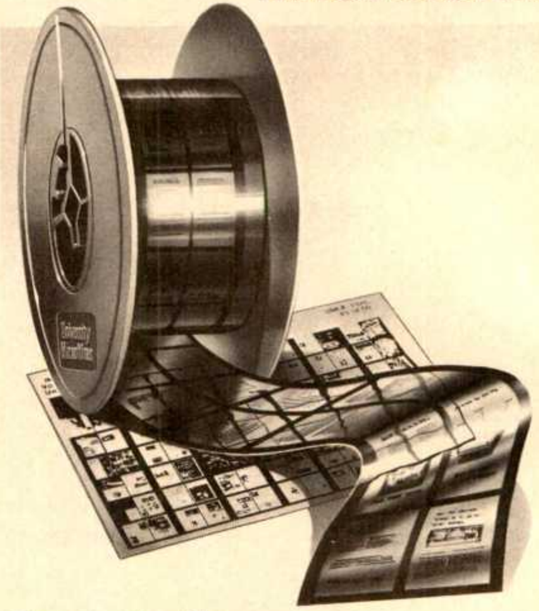
1 year(6 issues)for US$132.00(sea delivered)

Sole distributor **MARUZEN CO.LTD.** EXPORT DEPARTMENT
P.O.Box:5050. Tokyo International 100-31. Japan.
Telex:J26517 MARUZEN Cable:MARUYA Tokyo Fax:Tokyo 274-2270

Reader Service No.7

---

nature is available in microform.



University Microfilms International reproduces this publication in microform: microfiche and 16mm or 35mm film. For information about this publication or any of the more than 13,000 titles we offer, complete and mail the coupon to: University Microfilms International, 300 N. Zeeb Road, Ann Arbor, MI 48106. Call us toll-free for an immediate response: 800-521-3044. Or call collect in Michigan, Alaska and Hawaii: 313-761-4700.

☐ Please send information about these titles:

Name ____

Company/Institution ____

Address ____

City ____

State ____ Zip ____

Phone ( ) ____

University Microfilms International

Export market for advice

CAICYT is by and large a traditional research council: it assesses grant applications through the reports of external referees chosen by committees representing each subject area, it selects priority areas and arranges special grants to cover them and to aid cooperation with industry and the funding of research that might find application. All these activities and the other things it does (see opposite) are normal enough. What has made CAICYT a power in the land is that by all accounts it does its job very well and, more importantly, it was the first organization to bring an objective element into the assessment and planning of science. CAICYT's 300 or so referees have found themselves in demand. Grant proposals from both university and CSIC researchers normally pass before it (although if CAICYT chooses not to fund projects from CSIC researchers, they still have access to CSIC funds). And recently numerous other organizations have come to them for help, including some from abroad, like the Chilean research council.

The idea of making thorough analyses of the scientific situation really took off only recently. At the end of the Franco era one problem faced by both universities and the research council was that nobody really knew who was doing what or where and what equipment different laboratories had. Now CAICYT has a large computer and an enormous data base. A key word is sufficient to call up any scientist along with grant and publication record. Any piece of equipment can just as easily be located.

The data base has helped in putting forward realistic plans that Spain has the trained staff to meet, but CAICYT's concern for information sometimes appears to verge on the obsessional. The UNESCO coding system system (now, alas, out of date) is used and there is a joke in the laboratories that a researcher who cannot find a 6-digit category to fit his or her work will have a short career. ☐

that without a real commitment to spend more money (which the government has not made) no profound change can be expected. What is certainly clear is that the success of the new organization will depend largely on the quality of the advice it receives. Responsibility for that will rest heavily with a permanent expert commission set up to advise the Interministerial Commission. Fortunately, considerable experience is at hand in the form of a body that has already proven itself a driving force in the scientific scene — the Advisory Committee for Scientific and Technical Research (CAICYT) of the Ministry of Education. The new permanent commission will largely be the old CAICYT with a few additions and new name. ☐

The research council

Back on the road to recovery

THE Consejo Superior de Investigaciones Cientificas (CSIC), the nation's largest multidisciplinary research body, and its traditional source of scientific strength, is coming through a profound reorganization. At the end of the Franco era the CSIC had fallen so low that there was talk of abolishing it altogether. But instead it was reformed, more vigorously than the universities, and is reasserting itself both in basic and applied research.

The CSIC of the 1970s could provide a lesson in how not to run a research organization. The number of its institutes had grown to 170. Many of them were extremely small and had been set up for personal rather than scientific reasons; some contained just one person, for the old style Franco-era professor was a strong figure who found it hard to work under anyone else. Salaries were low and laboratories poorly equipped. The staff was growing older; by 1981 the average age was 48. But after the restoration of democracy things began to change and by the 1980s the pace had increased rapidly. Now just 76 institutes are left, some that vanished were closed down, but most were fused with others. That meant many former directors had, once more, to become subordinates or to leave.

Some were dispatched by lowering the age of compulsory retirement from 70 to 65. Not all went peacefully. Here and there in reorganized institutes there are autocrats who have hung onto large areas of laboratory space, despite the entreaties of their younger and more active colleagues. Elsewhere, quarrels over who was to gain power have left people no longer on speaking terms. And despite reorganization, integration of laboratories is sometimes proceeding in name only, but that perhaps requires only the passage of time to cure.

A flood of new appointments allowed young people to enter CSIC, or to move up its hierarchy. In 1981 there were 1,100 scientific staff; by the end of this year there will be 1,800. Over the same period the budget doubled and CSIC made efforts to open itself up to the universities and to industry. Fourteen joint centres have been set up in the universities, usually the university provides the building but all the laboratory equipment and around a third of the staff are provided by CSIC. Some 40 per cent of the Consejo's research projects also involve university researchers, for although they do not have direct access to CSIC's funds, they can make joint applications with CSIC researchers.

Industrial contracts are being signed at the rate of three a week. This year a little over a fifth of CSIC's total budget will come from industry. Industrial researchers can also come to CSIC laboratories for training — although it is not yet possible simply to send CSIC researchers to industry. CSIC's head, Enrique Trillas, is particularly proud of the successes of the Office for Evaluation and Technology Transfer and that industry is now coming to the CSIC without being asked.

RAMON y Cajal, who won the Nobel prize in 1906 along with Camille Golgi for establishing the neuron as the unit of the nervous system, signed, as its first president, the document that brought the Junta para Ampliacion de Estudiose Investigaciones Cientificas into existence in 1907. The Junta was to prove a liberal scientific institute sending many people abroad for training and leading a short scientific renaissance. Then the civil war came and in 1939 Franco suppressed the Junta. It was soon rebuilt as the CSIC; but not with the same people — some three-quarters of Spain's senior scientists had fled into exile. The new institution was given the quasi-religious goal of "rebuilding the unity of science broken by the philosophy of rationalism", but retained some liberal traditions; in the Franco regime's view, intellectuals were safer in the CSIC, where there were no students for them to teach, than in the universities.

Much then has been achieved. Even the Napoleonic style examination — the *oposcione* — has been altered, as it has in the universities (see p.320). Now, to get a first post within the research council, there is a public examination of the candidate's academic record and research speciality. But later promotion depends only on an assessment of the researcher's work. More important is that in both cases it is no longer a jury drawn from CSIC institutes only that makes the decision, but one containing professors from the universities and, increasingly, from abroad. In the past juries often came from the candidate's own institute, making the promotion of internal candidates and the exchange of favours a common feature of institute life. The rejuvenated CSIC institutes still span the whole of science: from history, through molecular biology to construction and cement.

For many in these laboratories, the greatest change has been that research proposals are now thoroughly assessed. In the old days a grant came with a permanent position. Now researchers receive a salary and basic infrastructure and must apply for grants, either from CAICYT and CSIC, private agencies or industry. And all scientific evaluations are made by CAICYT (see p.317) and its teams of referees, rather than CSIC itself, although an economic evaluation is first performed by a committee from CSIC.

CSIC's head Trillas.

The great changes made in CSIC do not mean, however, that its troubles are over. Although CSIC has better facilities for research than the universities and is the place where most research is done, it is still tiny, with less than 2,000 researchers, for a country the size of Spain. And more important, scientific traditions are not made overnight. The best institutes are at international levels. But they are few and far between compared to the run of the mill institutes that rarely produce a paper for an international journal. Moribund institutes still need infusions of talent, particularly of people trained abroad, but returnees face no easy task at institutes where excellence is not highly valued.

Most CSIC researchers, although happy to moan about the problems that remain, recognize the enormous improvements of the past ten years. If anything, what really needs to change is the attitudes of researchers themselves. One quote (delivered after a half hour of complaints about the bureaucracy) sums it up: "Our facilities are really no worse now than in the United Kingdom, and we have no lack of funds if we apply for them. It's up to us!"

Fundamental changes in attitude will come as talented young people take power within the institutes, coupled with increasing competition to gain grants. Then all that will be needed is a several-fold expansion of the system, before the research council begins to look like a ranking European scientific organization. □

Science evaluated

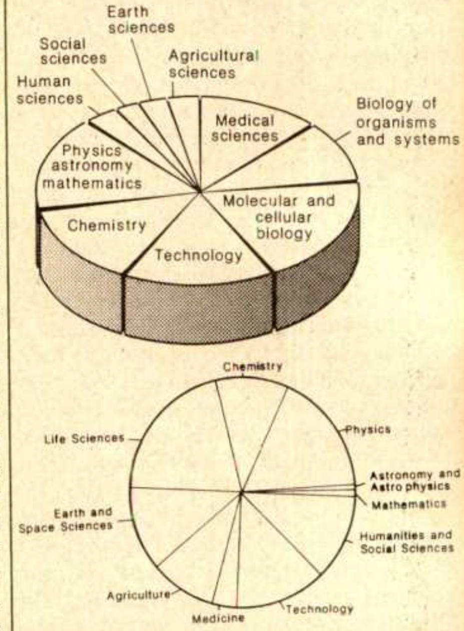
Still some catching up to do

It would be a mighty labour to try to give a complete and fair evaluation of the strengths and weakness of the different fields of science in Spain. Even in fields which are generally poorly developed there may still be outstanding characters making important international contributions: in ecology, for example, in which Spain produces few publications, there is Ramon Margalef at the University of Barcelona whose books have become standard works in quantitative ecology. Sweeping generalizations would not do such individuals justice. Something of Spain's record can, however, be read from the distribution of research grants of CAICYT and of research projects by CSIC coupled with the store of gossip available at every Spanish laboratory.

There is general agreement that molecular biology is strong (witness CBM, p. 319) and is attracting sizeable funding; "biochemists have had a powerful lobby for years" as one researcher put it. Physics has fared very much worse, with no large experimental facilities, and having seen people of the calibre of Blas Cabrera vanish into exile after the civil war. Only theoreticians managed to maintain standards. That should now change as a result of Spain rejoining CERN and an accompanying 'mobilization programme', in high-energy physics.

In nuclear physics there are some good Spanish scientists abroad but poor facilities within the country. European collaborations offer hope. Solid-state physics has also suffered from lack of equipment but this has already started to change.

Laser physics is the subject of a special CSIC mobilizing programme but a start is having to be made from almost nothing and standards are still low. In mathematics young returnees from foreign universities are just beginning to make things move and some specialized areas are already at international level. Astronomy appears to be a major disappointment. Despite access to the superb international facilities in the Canary Islands, very few Spanish papers appear in top journals. Atmospheric physics, earth sciences and whole organism biology remain very old-fashioned with only a few exceptions. But in geology, efforts are beginning to be made to link up with foreign researchers.



The share-out of research funds from CAICYT (top) and CSIC's project distribution.

Through English rags to scientific riches

Somewhat over half a century ago, Ramon Areces was working as a messenger boy in a little shop in Havana, Cuba. Now, through the foundation named after him, he is the biggest private source of support for scientific research in Spain. The path to philanthropy came through 'El Corte Ingles', a store he set up in Spain to make and sell suits with 'the English cut'. The name has stuck, although his department stores now sell everything from asparagus to zippers and it is impossible to find a city centre without one. And he is still selling suits — his factories producing ready-mades are the biggest in Europe. Overall, the group is one of Spain's biggest and parts of it have even gone into high technology — including electronics and biotechnology (see p. 325).

The foundation was set up ten years ago when Roman Areces, aged 72 and still a bachelor, decided to use part of his fortune to aid Spanish science. By international standards the foundation is not enormous but it is not so small either: In 1984–85, 147 million pesetas — close to a million dollars — were given out in grants along with double that amount for conferences, publications and to help equip research centres. Each year four research areas are chosen for support, and these replace (or are added to) the previously chosen areas. Anyone working in appropriate areas can apply for funding and grants are awarded on the decision of panels of referees. Research topics chosen recently range from robotics and solar energy to ocean sciences and cardiovascular disease. Individual grants are quite large — 23 million pesetas for a robot vision system, 13 million pesetas for molecular study of calcium channels in chromafin cells are examples. □

Centre for Molecular Biology

Long queues for laboratory space

"THE trouble with the Centro Biologia Molecular," said one research council administrator, "is that everyone wants to work there . . . but it's impossible". The 'CBM' has gained an international reputation for excellence and a national reputation for being far and away the best molecular biology research institute in Spain. Plus it has in residence Spain's Nobel prize winner Severo Ochoa. That has made it a popular place to be — a condition not without difficulties. Many at CBM are frustrated that they cannot accommodate all the researchers they would like while the research council would be happy to see some good scientists (particularly those returning from abroad) building good groups elsewhere rather than adding to CBM's lustre.

Among the main lines of research are virology; DNA regulation and protein synthesis; and molecular neurobiology, including the biochemistry of ageing and Alzheimer's disease. Perhaps best known (at least to readers of *Nature*) is the strong tradition in the developmental and molecular genetics of *Drosophila* begun by Antonio Garcia-Bellido on his return from Caltech and continued by Ginés Morata, Pedro Ripoll and Juan Modolell.

What makes the CBM interesting to the outsider is not just the science (of which there is plenty) but the way in which it has overcome some of the weaknesses in Spain's research system. It is the only major institute to have ended the traditional separation between CSIC institutes (specializing in research) and the universities (specializing in teaching). Under one roof (the university's) are the twin molecular biology institutes of the Autonomous University of Madrid and the CSIC.

The two sides share administration costs and elect their own representatives and a joint director. Those from the university side still have teaching duties while their CSIC colleagues have none, but that is the only difference and "we live in peace" as the director Jesus Avila put it. Both sides provide somewhat over 30 staff, and the CSIC over 60 'technical assistants'. Seventy-five research students complete the numbers. Fusion took place in 1975 before the need to bridge the gap between the universities and the research institutes was so widely recognized. But despite its success, the CBM has no imitators (only one institute has a similar structure). That might be because it took friends with political clout to set up the institute (among them Federico Mayor, then in high office). Their presence helps explain why there is no shortage of good equipment.

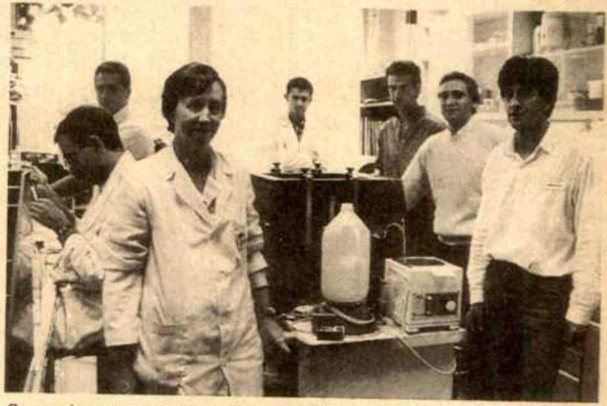
Severo Ochoa has now made his home at the CBM. He left Spain in 1936 and eventually became an American citizen working at the New York University College of Medicine. In 1959 he shared the Nobel prize with Arthur Kornberg for isolating enzymes that made it possible to synthesize RNA and DNA.

Organizational changes have helped overcome one common problem of the research institutes — a poor laboratory infrastructure (stores, animal rooms and the like). CBM researchers agree to apply jointly for a core grant as well as obtaining their own individual grants (although some researchers might end up with less this way) to maintain the research facilities. Elsewhere there is usually a levy on individual grants to provide common stocks but it rarely proves sufficient.

The core grant also enables them to tackle jointly one bizarre curse of CSIC laboratories: the technical staff belong to a large public service union and work from eight in the morning until three in the afternoon, so just when things are in full swing it is time to get the bus back to Madrid. Some extra money can be provided to keep some of them on later. And other problems, like keeping the heating on when the university switches it off (when classes are over) can be dealt with.

But despite these successes a gigantic problem remains to ire CBM researchers — the lack of space. There just is not room for everyone. "Asisnine" is how Dr Avila described the lack of coordination between the research council and the facility that has led to "at least eight people being employed who can't work here because they do not have space". A part of the problem is that space is not always distributed rationally. And that might be because, as a visiting advisory committee of foreign experts pointed out, the director has no real power. He or she is elected for two years which, according to Margarita Sallas, often seem "lost years from a scientific point of view, just a lot of administrative work". It is never easy to find someone to take on the job.

Without real power at the top, space cannot readily be allocated or projects terminated. Newcomers speak of having to bargain with territory holders to find a corner. As Juan Modolell put it "[space] is not dependent upon scientific accomplishment. I really have this laboratory because of historical circumstances, other people have half my space, or no space, and others several times more, but it has nothing to do with ability". That in turn, he says, reflects a more general problem, that "there is very little policy of promoting excellence".



Space is at a premium here in Margarita Sallas's laboratory as elsewhere in the CBM. The record seems to be ten people (and a few million fruit flies) in a room of just 49 square metres.

The director, Jesus Avila, does have plans to try to readjust laboratory space, if he is able (he was last seen by this correspondent wearily threatening to resign if promised administrative help did not materialize soon). What he wants to see is an institute with a potential for growth, that can let in young people. "Every institute has to grow or die, and we want this place to grow rationally", he says.

A more complete solution than readjustment can achieve might come with the new National Centre for Biotechnology to be constructed close by. There are doubts about just when it will be built (the decision was announced in 1984 but construction has not begun), but excellent molecular biologists are what it will need. There are many at the CBM who could do the job. A danger though is that the new institute will turn out to be just an annex to the CBM (or a place for some of its senior staff to retire to) instead of what it is meant to be — a laboratory working closely with industry. Those in industry who want something different from the CBM's strength in basic research complain that few people with industrial experience have been suggested for the board. For the researchers perhaps, as Juan Rojo, Secretary of State for the Universities and Research put if (very gently) "there may have to be a little shift of centre of gravity towards problems of interest to industry… a *Drosophila* group, if it's big, maybe some of them would shift a little to vaccines, agricultural products and so on." □

The universities

Worst problems overcome

Ask a group of professors about the state of the universities and you will hear a bewildering mixture of opinions. But after a little while a pattern emerges; praises come for the enormous improvements that have been wrought in the university system in the space of a few years but savage criticism is reserved for the defects that remain, often coupled with pessimism that they can ever be fully cured. Whether more criticism than praise is heard depends partly on the personality of the speakers, and very much on the amount of time spent abroad.

While perfection may be an impossible goal it would be too much to agree with the several cynics who said, "We were right at the bottom, so we could only get better". Planned reforms have had their effect and conditions for researchers are improving rapidly. The question is whether the reforms really went deep enough to give the momentum for the continued changes needed to bring Spanish universities up to the best international standard.

After the massive growth of the universities towards the end of Franco's rule real reform came with a new law for the universities — the *Ley de Reforma Universitaria*. Its overall aim was to remove the autocratic control exercised by central government over the universities, which made all universities equal and removed any spur to excellence.

Under the new law the universities were given power to write their own statutes, with the approval of the government, through the action of a University General Assembly (Claustro) freely elected by professors, students and administrative staff. Departments were rationalized. A huge number had been created under the old regime; many, it seems, to satisfy the personal ambitions of an individual professor or to separate those who had quarrelled. Resources are now concentrated in single departments for each subject area.

The enormous problem of untenured teachers was also tackled. The rapid rise in student numbers during the 1970s had made it impossible for the universities to appoint sufficient staff through the usual bureaucratic channels and many lecturers were appointed on short-term contracts. By the end of the boom there were more of them than there were tenured professors. Commissions were set up under the new law to decide who among the non-tenured staff should be taken on as professors; not surprisingly, perhaps, difficult choices were often avoided and the great majority were found suitable.

With their own charters written and

Research conditions still lagging behind

The University of Barcelona's School of Biological Science boasts an impressive piece of modern architecture, full of complex curves on the outside and with a huge central well inside, around which stairs zigzag upwards. But the building was planned in a more autocratic era and little consideration was given to the needs of those who were going to use it — with the peculiar result that corridors pass through the middle of research laboratories and libraries while unusable empty spaces abound.

Space is one thing the school is short of, for the study of biology attracts vast numbers of students; 3,000 in all. With numbers like that, some of the greatest difficulties faced by the university teaching system can be clearly seen at the school. There is no selection of students entering the school (biology is often a second option for those who wanted to study medicine, one of the few courses where numbers can be limited) and after one year only 800 of the 1,500 who entered are left — a drop-out ratio also seen in many other places.

At the end of the five-year course, half of the graduates will be unable to find a job, and only 10 per cent of them a job related to biology. With these prospects, it is not surprising that just half the students complete the course in five years; others retake courses they failed *en route* a few times — some stay on ten years and are well into their 30s. Only 10–20 per cent of the students have grants so the prospect of money running out does not provide a spur.

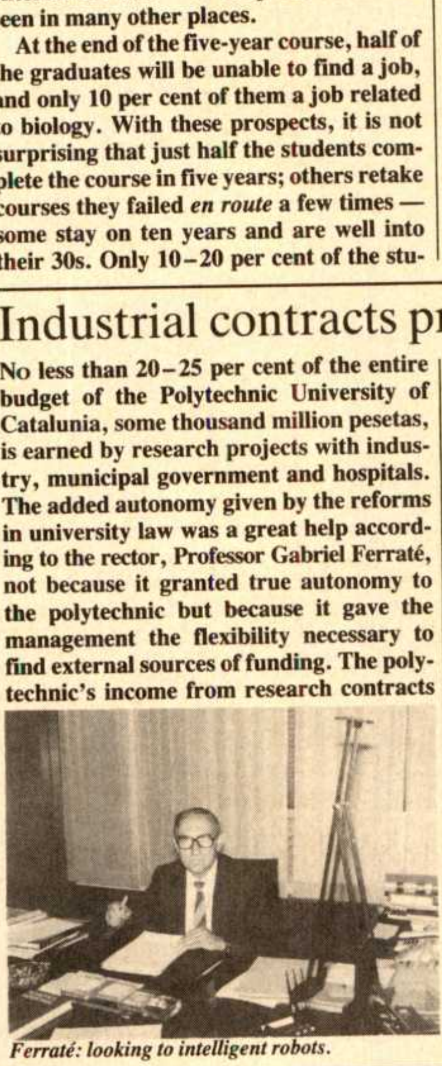
At the start, classes are big (more than a hundred students in each) and there is no chance of tutorials or seminars, nor much practical teaching — making the courses almost 'theoretical biology'. Only in the later years can groups of 20–30 students be dealt with. Given these factors, few students are highly motivated, and the long teaching hours make a real research effort from the staff difficult to attain. On top of that, back-up services are almost nonexistent.

Impressive architecture but does it work?

Professor Jaume Baguñá, who heads the Department of Genetics, has a dozen researchers and the same number of doctoral students without a single technician and only a part-time administrator. As Professor Baguñá says, "things have improved and change is coming, very slowly. But university research conditions are still very bad compared to the rest of Europe". □

Industrial contracts provide rich harvest

No less than 20–25 per cent of the entire budget of the Polytechnic University of Catalunia, some thousand million pesetas, is earned by research projects with industry, municipal government and hospitals. The added autonomy given by the reforms in university law was a great help according to the rector, Professor Gabriel Ferraté, not because it granted true autonomy to the polytechnic but because it gave the management the flexibility necessary to find external sources of funding. The polytechnic's income from research contracts is now increasing almost exponentially.



Ferraté: looking to intelligent robots.

The polytechnic was one of those in the camp opposed to keeping the status of teachers as civil servants, according to Professor Ferraté. But under the new law, winning contracts can also mean an increase in salary and that helps to provide motivation, which the polytechnic is taking full advantage of.

Research at the polytechnic has been buoyed up by contracts with young and energetic Catalan firms in the new technologies. Although the school of civil engineering has a high reputation, real glamour, as elsewhere, has passed to computer sciences, computer-assisted design and manufacturing and telecommunications: successful projects have been run on fast pattern recognition systems and automatic map-making from aerial photographs. Professor Ferraté himself founded a hybrid Institute of Cybernetics combining polytechnic and CSIC facilities. Its main thrust is "artificial intelligence applied to robotics", and several cooperative projects are running with US universities. □

approved what will happen to the universities? Will each university acquire its own distinct character rather than being one unit of a centralized bureaucracy? Will the best researchers (or teachers) find themselves in demand from competing universities? A glance at the law suggests that the answer will not be simple. It is a remarkably large document — some 59 articles and 22 pages. The autonomy given to the universities is of a form which it prescribes very exactly. All are now autonomous — but in the same way.

A key problem, not fully resolved, is that it is hard to reward directly those who are talented. There are only two full-time positions; the full professor (*Cathedraticos*) and assistant professors (*Profesores Titulares*). Salaries are fixed. There are no grades, and there are no easy means to increase a professor's salary (other than the fixed annual increment). This has meant there has been no means whatever to try to attract the brightest and best researchers, nor any reward for creating excellence. The result, over the years, has been an almighty indifference to the quality of research and teaching at a university; as one young researcher at the Barcelona Autonoma put it, "no-one really cares if their university is good or not. No-one ever fights to get the best man. They are happier with someone worse who won't provide any competition". And without any incentive to find good people the tendency has been for those already in place simply to rise through the ranks. The key step to a successful career has always been regarded as forming a close attachment to a powerful professor early on. The new law does allow researchers to double their salaries through contracts with industry, an incentive that many universities, particularly the polytechnics, are grasping with vigour. But the law provides little stimulus for basic research.

While autonomy may not have gone far enough in salaries, many say it has gone too far on appointments. The extra freedom given to the universities to decide on their staff has not always been wisely used — although it is still far from complete freedom.

All university staff are civil servants and in the past appointments were centrally controlled. Each required one of Spain's great official rituals — the *oposicione*. In a series of public examinations candidates had to prove their worth, sometimes in a direct debate with their competitors, before a tribunal of seven appointed by the government. The range of questions that might be asked was huge; it was not uncommon for candidates to be given a topic and be required to deliver a lecture on it the next morning. Nowadays, things are much simpler and the procedure is closer to a formal panel interview than a public ordeal. The examination comes in two parts; the first on the candidate's curricu-

Comfortable surroundings and a strong Jesuit teaching tradition make Comillas popular.

Comillas can set high entry standards

One advantage of being a private university is that you can select your students — and to enter the Universidad Pontificia Comillas, run by the Jesuits, is by no means easy. Exam grades have to be very high and are followed up by psychological tests and an interview. Part of the reason for the university's success is that its students almost always find jobs. "People trust our recommendation", says the university's rector Guillermo Rodriguez-Izquierdo. Many applicants are the children of graduates; it is apparently a painful process to turn away those of them that are not up to the mark academically.

The university, now in Madrid, began as a seminary set-up at the end of the last century and its degrees are recognized by the state following an agreement between the Holy See and the Spanish government. Business administration and economics are now the most sought-after courses for the university's 6,000 students, but there is a popular course in engineering and a unique system by which industry supports doctoral students.

The university is aiming to achieve a different style from the others. "We try to combine science with radical concern for human problems", says the rector. And perhaps just as important, "students come to the classes on time!" □

Getting basic research on the rails

The Autonomous University of Barcelona must be the only university in Spain to have its own purpose-built railway terminus. That is because the university was built, in 1968, 20 kilometres outside Barcelona with nowhere nearby for its 28,000 students to live. The university's newly elected rector, Professor Ramón Pasqual, readily admits that this has put the university at a disadvantage and that good students may choose to go elsewhere. But he is one of the optimists about the university system's future, and of the ability of his own university to offer something special. "Under the new law," he says, "the most important thing is that you can be different. In a few years there will be good and bad universities and the Ministry of Education will play differently with different universities. Some will go well, others will have problems".

New joint CSIC university institutes, including those for microelectronics and for materials are being set up on the campus and a science park is appearing across the road (Catalonia's 'Silicon Valley'). Links with industry are growing.

Pasqual is also optimistic about Spain's basic research effort. "Now every journal has a paper from a Spanish researcher", he says. And as a theoretical physicist he applauds the decision to join CERN and the European Synchrotron Radiation Facility which will provide completely new opportunities for experimentalists. The big problems that remain? "Universities are still very poor. We lack technicians and we need administrative support — departments here still have no secretaries". □

A 'campus university' in the true sense.

lum vitae, and the second on the field of research. The experience can still be harrowing, as much for the tribunal as for the candidate. When there are many applicants each has to pass through the same lengthy procedure and a tribunal can be busy for a month.

Where the university has gained control is in the selection of the tribunal. Two of the five members are provided by the university that has the vacancy; the remainder are chosen at random from among professors in the same field. At the same time the university may provide a clear 'profile' of the person they want to fill the job; the combined result is that it has become much easier for the universities to choose who they really want. And what it seems they have all too often wanted is simply to promote someone in the same department. Greater autonomy has, if anything, increased inbreeding rather than fluidity among researchers.

There are even stories that universities have connived to avoid a regulation that says no one should be appointed who has not spent at least some time at another university by making linked appointments in which professors apparently move but in fact stay where they are. And these problems come at a time when there has been a boom in new posts.

Just how bad an effect this has had is very much a matter of opinion; some go as far as to say that the net result of having to give many untenured teachers tenure, alongside misuse of the selection system, has filled the universities with mediocre young professors who will now dominate university life for thirty years. Others say these are just teething problems. And some politicians blame the rectors for the universities' problems. Rectors are elected under the new law and this makes them anxious to please their constituents rather than make strong decisions. Certainly, many in the political world now regret that rectors are not appointed.

Reform could have been carried much further by doing away with the automatic appointment of staff as civil servants with jobs for life. There were many, including very senior staff, who fought to replace lifetime tenure with long-term contracts to enable selection to be repeated at intervals. Curiously, the socialist government was of the same mind, before it faced the challenge of putting its ideas into practice.

One way of quickly raising standards of research is to send people off to the world's very best institutes for several years of doctoral or postdoctoral training. Spain has certainly not ignored this option and several exchange agreements and fellowship schemes exist. But those who go abroad say their problems begin when they try to come back. Foreign degrees are not automatically recognized, nor can a returning scientist expect an automatic right to the same amount of laboratory space and access to equipment as those who have built their territories at home. The administration is aware of the problem — and some who have returned successfully say that those who have not do protest too much. All want to work in the very top laboratories, ignoring the importance of establishing groups in the smaller universities.

One part of the university reform that seems to have had little impact is the decision to transfer responsibility for the universities to the autonomous regions. In practice this has not meant universities acquiring a strong regional flavour. All that has happened is that administration has been made somewhat easier, for not all requests have to go as far as Madrid. But Madrid still holds the budget strings and transfer has not been total. Half of the universities (including, naturally, those based in Madrid) have simply chosen to leave things as they are. And the grass always seeming greener on the other side of the fence, those universities that have been transferred protest that Madrid universities are able to collect more money, while the Madrid universities claim the others collect funds both from the central and the regional government. □

Complutense tackles the problems of scale

The Complutense in Madrid, with 118,000 students and 5,000 professors in 18 faculties, is one of the biggest universities in the world and by far the largest in Spain. One in five of all students are educated here. Managing a university on this scale is "often terrible; each small problem is multiplied by hundreds", says the vice-rector, Professor Angel Martin Municio.

In research, new measures are being taken to break down the barriers between disciplines and between the many small empires which have traditionally formed in the universities. The carrot being offered to persuade groups to get together is expensive technology: a world class nuclear magnetic resonance imaging centre is being set up to serve all researchers in the university, for example. Another success has been in bringing together all the relevant groups in the six hospitals attached to the medical school by providing funds for animals for use as disease models. This way the university also gains facilities good enough to allow participation in international programmes.

Headway also seems to be being made with the problem of student wastage; some departments in some universities lose half their students after one year. This year, within Madrid's three universities, students applied for a subject rather than a university. The result is that 80–90 per cent are studying what they want and motivation has been improved.

As elsewhere, the Complutense is actively pursuing contacts with industry. In five or six years, Professor Martin says, "Universities will be different according to their courage and effort". □

Professor Martin needs a helping hand.

Autonoma tops for molecular biology

The campus of the Autonoma in Madrid is instantly recognized as 1960s; bits of decaying abstract sculpture dot a park surrounding concrete buildings connected with tunnels, perforated by underpasses and decorated with murals symbolizing something that has long passed out of common knowledge. Just as at an airport you know what kind of place you are in — even if you are not sure of the continent (except here, perhaps, for the rare fragrance from the surrounding olive groves). And like an airport, the biggest facility is the car park; the campus is a half-hour ride from central Madrid and there seems to be an endless line of students trying to hitchhike a ride back there.

But being close to Madrid has helped the university attract some of the country's best researchers. The Centre for Molecular Biology is far and away the top institute in its field and there are strong groups in chemistry and in physics. The new Institute for Biotechnology is to find its home here. □

Molecular biology and sculpture predominate.

Science policy

The man in the hot seat

Professor Juan Rojo, a physicist, is Secretary of State for Universities and Research, which gives him overall responsibility for both the university system and CSIC. It is a job which he describes as "bad in the sense that I have to deal with personnel problems but good in that I can contribute to making stronger links between the universities and the research system". Professor Rojo's horizons will soon widen more than a little: once the new Interministerial Commission is running he will lead its advisory committee and that will give him a chance to look at the government's total 100 thousand million pesetas spent on science.

Professor Rojo seems optimistic about the universities' future. To those who believe that giving increased autonomy to the universities has actually increased their "inbred nature" he says. "Some, it is true, have used their autonomy to promote their own people but if you give autonomy you cannot control the universities at the same time. I believe autonomy will open the door for excellence." He is proud of another aspect of the law of university reform that is not widely commented on: the creation of the "Social Council", which contains 12 people from outside the university (including representatives of industry and the trade unions) and eight from inside. This "provides a counterbalance to autonomy", and he says "we have good evidence that it is fighting inbreeding". The council has powers "to increase the salaries of professors of special excellence, and to decide whether vacant chairs should be transferred to new areas . . . or, say into two lectureships".

Professor Juan Rojo want stronger links between universities and the research system.

Professor Rojo is also tackling another source of complaint: that Spanish researchers who have worked successfully abroad find bureaucratic obstacles prevent them returning easily. First task is to make a standard list of centres from which degrees will be recognized, "so that people don't have to waste time translating their theses" (for doctoral degrees and so on to be locally validated). And "second, we are trying to favour the kinetics of return. We are encouraging people to take their sabbatical leave here, as well as giving 2-year fellowships".

On a wider issue, work for the new Interministerial Commission has already begun. Themes under discussion for national programmes are information technology, biotechnology, new materials, engineering sciences and food technology.

Although many feel that Spain's total research expenditure is shockingly low at 0.6–0.7 per cent of gross national product. Professor Rojo points out that the rate of growth has been high, coming up from 0.4 per cent in the past five years. The top priority now is "investing in the training of people for the future. We now have 4,500 people at the postgraduate level which is reasonable". But there is still a serious gap to be filled in the training of technicians. "For people of 18 there is very little alternative to going to university, there are practically no technical colleges at all and we need to develop the level of technical training." ☐

Agricultural research

Molecular biology in the pig-pen

Agriculture is the only sector in which most research centres have been transferred to the autonomous communities. A precedent for a fuller devolution of research was not set, it seems, because agriculture is recognizably different. Each area has its own agricultural problems and needs its own research institutes to tackle them. Their activities are, however, closely coordinated from the centre by the National Institute for Agricultural Research (INIA).

Under the new Law of Science programmes for agriculture remain the responsibility of the Minister of Agriculture, but they are complemented by more wide-ranging interdisciplinary programmes that include other ministries and are parts of the national plan. Programmes are monitored by a commission from INIA containing referees from independent organizations.

One research institute, in Madrid, is directly attached to INIA (and takes its name from it) and is the home of most of agriculture's high-level basic research, including work on one of Spain's biggest agricultural problems, African swine fever. Spain and Portugal are the only countries in Europe where the disease is widespread. That means that although pigs are Spain's biggest agricultural earner (twelve million head a year are consumed), exports are almost entirely prohibited for the disease spreads easily, particularly in contaminated pork, and is usually fatal.

It was the rise of air travel that brought the disease to Iberia. Pigs kept near Lisbon airport, fed on swill from Angola, were the first to be infected. From there the disease spread into Spain in the early 1960s, before much was known about it. Mass slaughter of infected animals has proved the only method of control. When the disease reached the Dominican Republic in the 1970s the drastic decision was taken to slaughter all the pigs in the country. Small outbreaks in North Europe have been contained by the same means.

In Spain slaughter of infected animals and strict control of movement of stock mean that eradication may be possible in a few years. But now the infection has spread into the wild population, periodic re-infection is probable. The only answer, it seems, is an effective vaccine. But the search for such a vaccine has often seemed a hopeless task for the virus does not stimulate production of neutralizing antibody, a property shared by one rare mammalian virus; that causing Aleutian disease in mink.

Work at the institute now concentrates on understanding the basic molecular biology of the virus, immunological reactions to it, and its infectivity in pork products. Study of the virus genome does not encourage hopes that a vaccine can be produced easily. A variable region containing several gene families has been identified which suggests that the virus escapes the immune system by constantly changing. Now it is important to find out exactly what the genes code for.

The immunologists are a little more hopeful. Some pigs do recover from the disease and overall mortality has decreased over the years, suggesting that the immune system can do something. Some experiments suggest that the virus escapes from the host's immune system not because it is changing but because viral proteins are masked by host proteins. There are now hopes that there are some small virus-specific proteins that are not masked and against which antibodies can be produced. Serum from pigs that recovered from infection has already been found to inhibit infection by the virus *in vitro* tests. Can this encouraging result be built on in the future?

What the researchers now need ("worth their weight in gold", according to the institute's director Dr J. Sanchez-Vizcaino) is pigs that are serologically positive but healthy. That and another couple of years work will show if there is any hope of a vaccine. Research will soon be intensified with a move, 40km from Madrid, to a new Centre for Animal Infection, one of the largest in Europe. It will be perhaps the only one to have a large part of its floor space dedicated to pig-pens along with P3 and P4 laboratories. ☐

Industrial research

Companies need more science

SPAIN's entry into the EEC, on 1 January 1986, was the ultimate proof that the introspective Spain of the Franco era was dead. But enthusiasm for the wider horizons of Europe has to be balanced with the economic realities — Spain's industry now faces its biggest challenge for decades.

The key problem is that industry has not invested in research and development and that has left Spain with a pitifully small total research expenditure — just 0.6 per cent of gross national product, against the 2.4 per cent of Britain and the 2.7 per cent of West Germany — and largely low-tech industries. Boosting the research figures is now one of the government's top priorities.

Every research body is busy seeking out links with industry. The research council (CSIC) can boast that it is signing three contracts a week with industry on behalf of its institutes: and the polytechnic universities are doing particularly well with income from industry rising exponentially in some. Researchers are clearly under pressure to push industry, but is industry really prepared to pull? Dr Joan Majó, until a few months ago Minister of Industry and Energy and an electrical engineer, gives his views.

Former industry minister Joan Majó sees Spain as a major food-producer in years to come.

"There is not enough research in industry, that is clear. People simply blame the industrialists for this but it is more the fault of the circumstances. For years we have been living in a closed market, without external competition and there has been no need for R&D to make money. Some companies with external trade recognize the need and have reached good technological levels. But we do not have any very big multinationals that can support huge research efforts. Instead we have many medium-sized companies."

Can industry survive EEC entry and reach European levels? "We are joining at the best moment in the last ten years. Why? Because the opportunities are now greater than the dangers. Five years ago it was different. Industry had still not reacted to the industrial crisis of the 1970s and there was overproduction in Europe for which Spain could have provided a new outlet. Change began in the early 1980s and now we are able to compete. Productivity and investment have risen although some sectors, like iron and steel, have special problems."

Dr Majó believes that in any case Spain does not need to do all its own research. "The problems of the use and generation and technology are not the same. We have to accept that part of the technology we need is not generated here, or is developed by external firms. External firms should be encouraged where Spanish ones have no chance of success in the short term. The key is to be selective."

A range of familiar incentives — tax breaks, import restrictions and direct grants for research — are being used to stimulate industrial research. Some 35,000 million pesetas went in direct grants to electronics companies during Dr Majó's own period in office from 1982 to 1986. But there is still a lack of research infrastructure and trained personnel. One danger is "a possible collapse of the supply of technicians caused by the small scale of the technical universities. If major industries did all set up research laboratories they would exhaust the supply of trained personnel."

Top priorities are the same as elsewhere: electronics and biotechnology. "But Spain should become one of Europe's biggest food producers in the next few years and the biotechnology to be sought is to improve processes in food treatment rather than new pharmaceuticals".

Attempts are being made to build bridges between the universities, research institutes and industry, but how successful are they?" Efforts beyond the new Law of Science will be required. I have supported giving preference in grants to industry if they are collaborating with the universities, and to those who have joined multinational projects. And the technical universities need more resources". □

Industrial research

Looking for high-tech winners

THE basic research of yesterday is the applied research of today and the product of tomorrow. Most governments already realize that special mechanisms are needed to help speed the transformation of ideas into realities if they are not to be left behind in the international high-technology race. In Spain the need is perhaps more acute than elsewhere, for the banking and industrial communities are conservative and show little taste for risky ventures. So the government has provided the Centre for the Development of Industrial Technology (CDTI).

CDTI was set up back in 1977 but it seems to have only really got cracking after it was set free in 1982. It now has the status of a private company, although its budget and its board of directors are still provided by the government. Its director, Dr Jaime Sodupe, sees it as an "interface between industry and science" but talking more to the industry side as "there are already plenty of institutions to worry about science". Its activities carry on where CAICYT leaves off; support to help develop a small (or sometimes not so small) company's prototype so that it can enter the marketplace comes a stage after the applied research backed by CAICYT.

Some 600,000 million pesetas are now pumped by CDTI's 50 staff into its 100–200 projects each year. A part of the money is given out in venture capital repayable after two to six years. But the important point is that repayment is linked to commercial success; if the product fails CDTI can say goodbye to its money. Products already close to the market are often financed with an unsecured loan — without a need for guarantees the company's ability to borrow further from the banks is not affected. So far CDTI has had more than 60 per cent of its money returned safely which makes the director wonder if they are being sufficiently daring; similar organizations elsewhere in Europe often obtain only 30 per cent return.

Somewhat unusual is CDTI's role in international affairs. It is responsible for managing Spain's annual contribution to CERN (2,000 million pesetas), the European Space Agency (5,000 million pesetas) and soon the Airbus. And it also tries to find ways to get companies into European programmes. There have been successes. At the Eureka meeting in London in June, 67 projects were involved; 14 of them involving Spain, six of these with Spain as chief partner. CDTI's staff is optimistic, "each month is easier than the last, we don't have to convince managers anymore", said one of them. One remaining problem is that a big part of CDTI's job is also to collect and distribute information, and make industry aware of the opportunities and the latest world-wide trends (as well as organizing promotion of Spanish industry at fairs and the like). But Spain's embassies abroad are still living in the past — they have no science attachés. □

Industry: biotechnology

A sunny spot for entrepreneurs

SPAIN has not been immune from the world-wide rush towards the commercial exploitation of the advances of molecular biology. As elsewhere, the United States has provided the inspiration for scientists to turn entrepreneur — often directly, for the staff of the new biotechnology companies mostly have excellent American accents acquired at top US universities. The risk capital behind them, though, is Spanish: often it is CDTI (opposite) that is playing fairy godmother.

Five small companies are in the market for diagnostic kits using monoclonal antibodies. Among them are the well-established Biokit, now participating in a Eureka project to develop a gonorrhoea diagnosis kit, and the newer Invesgen which among other things is looking for a test for heroin consumption (proving between them that there is money for molecular biology in moral decay).

Biokit, described by a competitor as "a very pushy lot", can boast that the average age of their staff is just 30 years. They were the first to manufacture diagnostic kits in Spain, with a test they produced for rheumatoid factor back in 1976, just two years after setting up their company with ten people in a "120 square metre space". World-wide marketing agreements with foreign companies followed for this and for their next product, a syphillis diagnostic. With financial help from CDTI, chlamydia, rubella and hepatitis diagnostics came next. The philosophy of Biokit when aiming at a new product is "we must have contact with everyone in the world with knowledge in that field". What this means in practice is that despite being based in Barcelona, Biokit scientists go out to the United States knocking on people's doors until they are known and trusted and it is understood that the little company from Spain should be taken seriously.

The future now looks bright. Many of the products under research will be launched in the next year or two and there is hope that sales will reach the 1,000 million peseta mark and Biokit be a world leader in sexual disease diagnostic kits. But still the major input of expertise is likely to be US not Spanish universities. For the distant future (which is more than three years ahead in this business) there is hope from the Eureka project. But the race to develop a quick test for gonorrhea is not going to be an easy one to win.

Invesgen has not had to work so hard in building links to the United States as its key personnel came from Massachusetts Institute of Technology (MIT). Indeed, Invesgen probably only exists because the government decided back in 1974 to send thirty graduates to MIT to study for their doctorates. On their return some were involved in a project to set up Spain's own miniature MIT — an idea that eventually proved to be ahead of its time. Two years ago, four of the original MIT students got together with a fifth researcher to found Invesgen, at first within a division of a huge company of the El Corte Inglés group and soon after as an independent company in the suburbs of Madrid.

A start in setting up a sales network came through an agreement to market the Pasteur Institute's products; and a push for new products came with finance from CDTI. The company has grown to employ ten postdoctoral researchers plus technicians and sales staff. Closest to the market is a kit to quantify simply factors related to rheumatic disease (C-reactive protein, streptolysin, and rheumatoid factor). Selling Spanish products to hospitals in Spain is not going to be easy for there is still a feeling that foreign products are better. But the company's research director, Dr Julio Coll Morales, quotes the adage "no-one profits in his own land" and says that progress will come through selling in the United States. Distance from the world's scientific giants does make progress harder; at the most direct level there can be very long waits for reagents, sometimes so long that "people forget why they ordered them by the time they arrive".

And Dr Coll Morales admits that they are "a little lonely" working in a small group far from the universities. There is still some predjudice against the "fellow from industry". Hopes are that the new Centre for Biotechnology will make a difference. □

Ports safe in a storm

PROFESSOR Felipe Martínez and Dr Manuel Pastor of the Research Centre for Public Works (CEDEX) stand in front of Bilbao harbour. Not the real thing, but a massive model set up in a gigantic building complete with wave-making machines. Under the same roof are the harbours of Valencia, Las Palmas, and part of Barcelona beach needing restoring before the Olympics. The Centre for the Study of Ports and Coasts, where Professor Martinez is director, looks at all kinds of problems in coastal defence and harbour construction: particularly important as Spain has one of the world's largest fishing fleets, much of it operating from small ports. At Bilbao the problems are for big ships: a peculiar resonance in the harbour at times amplifies long-period waves (of four minute period and unknown origin) and makes unloading of ships at one quay difficult. Dr Pastor's finite element model and the physical model itself both produce the same effect and suggest a cure — opening an extra narrow channel between two basins. □

Patent success for pharmaceutical group

FERRER International is an exception to the rule that Spain's companies do not invest in research and technology. The company is not among Spain's 200 biggest enterprises — but when it comes to total research and development expenditure it ranks twelfth. And its researchers boast more patents than anyone else in the private sector. Its business is pharmaceuticals.

Change in the company's thinking — a determination to make new rather than derivative products — came in the 1970s alongside the nation's political liberalization. An effort was made to bring research up to best European levels and now over 10 per cent of Ferrer's 1,000 employees work in its Barcelona research centre.

For the centre's director, there is no choice for Spain but to try to increase its technological level, "it's either technology or colonization", he says. But the path ahead is by no means easy; there are plenty of giants in the pharmaceutical business and Ferrer International is just medium-sized, with annual sales worth some 15,000 million pesetas. To succeed, the company will have to rely on "agility" and sticking to fields where they have good ideas, mainly analgesics, antidepressants, Ca^{2+} antagonists, H_2 histamine receptor blockers and antifungal drugs.

Their approach is conventional: studying physiology and the structure of chemicals known to have physiological effects modifying and testing them and improving routes of chemical synthesis. Collaborations with universities are common: "universities are changing", says the director, "but we must be realistic. It is difficult for them. They often do not have the best apparatus and they must think of how many publications they can produce. Industrial interest is not yet sufficient to resolve their problems." □

The Basque Country

A mountain of problems to tackle

THE Basque Country suffered from cruel neglect under the Franco regime. Its heartland, the provinces of Guipuzcoa and Biscay (which contains the town of Guernica), had fought on the side of the Republic during the civil war and been granted autonomy by it. With Franco's victory autonomy ended; the region was 'punished' and use of the Basque language was banned. Despite the region's enormous industrial importance the central government refused to establish a university or research institute there — just a school of engineering. Now the technological backup to help replace the region's declining heavy industry is just being developed. And out of hardship there has grown a unique Basque response in indigenous cooperatives that link industries, research institutes and a bank.

Parts of Bilbao, where heavy industry was most concentrated, are palpably dying. The town grew rich in the First World War when Spain was neutral and both sides were desperate to buy its iron and steel. Several of its biggest works are now bankrupt but the city, trapped in a narrow valley, still spends half its time blanketed in smog. Signs of decay are everywhere; the unemployed crowd street corners and the walls are covered with slogans, both political (in Basque) and the more desperate "*Heroina Assassasina*". The worst blows, occasioned by entry into the EEC following on a world recession, are thought to be over. The contraction in the steel- and ship-yards is coming to a halt but a large number of small and medium-sized manufacturing businesses remain that have never had the money to invest in research and whose equipment is out of date. And in a story familiar from Sunderland to Pittsburgh, the race is on to add value: control and automation, computer-assisted design and manufacturing, robotics and flexible manufacturing are the key words.

There is only one area that is taboo: nuclear power. Construction of a nuclear power plant was begun but abandoned after the project manager, and then his successor, were assassinated by Basque Nationalists. Although nationalist sentiments have lessened with regional autonomy, terrorism has still to be taken very seriously.

The regional government is doing what it can for science and technology within its limited means. It has taken over responsibility for several research centres intended to raise industry's capabilities. Typical of them is Labeine in Bilbao. Set up in 1955, when research was more a "desire than a reality", it has expanded at a cracking pace (40 per cent a year) since the local government gained responsibility in 1980. Its budget is now 600 million pesetas (£3 million), still peanuts by international standards; and its staff 130, half of them scientists. The intention is not to do basic research but help industry; both at the level of testing and pollution control and the introduction of high technology. Among their projects are those for developing an expert system to diagnose faults in electric power stations and to establish a unit for the supply of customized integrated circuits.

A new university has been established, scattered over the region but with most of its departments in Bilbao and San Sebastian. It began life in 1968 as the University of Bilbao; addition of other centres turned it into the University of the Basque Country, now with a little over 40,000 students. The quality of the university remains very uneven and it proves hard to recruit good people in those areas, like computing and engineering, where they are most needed, for industry can always offer better pay. A massive shortfall of qualified people in these areas is forecast for the end of the century, not so much through lack of students but through lack of competent teachers.

Basic science is being provided for in scholarship schemes that are unique to the Basque region, but which the central government could do well to emulate. One hundred and fifty predoctoral and twenty postdoctoral grants are being given (held for two years in Spain, or three years in a foreign country) to help raise young researchers' standards. About half the grant holders stay in the country; for the rest France, the United Kingdom and then the United States are favourite destinations (and cost the government only a little over a million pesetas a year). Attempts are also being made to bring the university up to international level by providing money for re-equipping laboratories. Four hundred million pesetas were spent in this way last year, but applications were for four times that amount and the search is on for new sources of funds. Industry is more directly supported through spending 1,000 million pesetas a year on projects in the regions' five industrial research centres (including Labein and Ikerlan described here) to develop 'generic technologies' to the level where industry can come in for specific applications.

All this support comes from local taxes, quite separately from payments made to the central government for 'national' services — which still includes science and technology. Some of the money comes back in grants from CAICYT but 70 per cent of the nation's research institutes are found in Madrid. Not surprisingly there are local politicians who strongly feel that the contribution for the science budget should be handed back to the Basque government to spend as it wishes in its own area.

The difference of opinion stems from the failure of the constitution to be crystal clear on the dividing line between the competence of the autonomous regions and the central government. Does the power of the government to coordinate research mean that it also says how the money is to be spent? The new law of science, which increases central control, raises hackles in the Basque Country and many declare it unconstitutional. More rumblings of discontent from them (and from Catalan politicians) are bound to follow but it would take a narrowly won election, with nationalist parties holding the balance, to persuade the central government, whatever the constitution may say. □

Mondragon, amid mountains in the heart of the Basque Country, is packed with industrial cooperatives.

The Basque Country

Cooperatives provide a solution

ANYONE who thinks that cooperatives, in which workers own their workplaces, could never be more than an idealist's dream should visit Mondragon in the Basque Country. Britain's Minister of Employment, Kenneth Clark, made the effort last month and apparently came away impressed. Mondragon's industrial cooperatives now have 19,000 employees (or employers, depending on which way you look at it) and their export sales are booming. Alongside them are agricultural, education, consumer, housing and services cooperatives which can provide everything from technical education to part-time jobs for women with children. And the industrial cooperatives *en masse* have felt the need to raise their technological standards sufficient to establish Ikerlan — a unique cooperative research institute.

The cooperative movement grew in the special conditions of the Basque Country from the vision of one extraordinary man, Father Don José María Arizmendiarrieta. During the civil war he worked as a Basque Army journalist and at its end he narrowly escaped execution. He came to Mondragon, a small town in the mountains south of Bilbao, and in 1943 set up a small technical training school with the aim of aiding the reconstruction of his country. Those enrolled not only studied but raised money from the community to keep the school going. As it turned out, the graduates of the school were later to become the managers and technicians of the cooperatives. The first was established in 1956, after a vain attempt to persuade local industry to allow worker participation in ownership. It consisted of just five men making paraffin heaters. Within a year there were five more cooperatives. They were a success; perhaps partly because the Basque Country has a long history of communal labour on its farms. Then the problem became how to raise capital for further expansion. The answer: set up a cooperative bank. More than twenty years later the Caja Laboral Popular (the working people's savings bank) has over 120,000 million pesetas (£600 million) on deposit and provides financial management, export advice and banking services to the region's 160 plus cooperatives. Its funds have come from the local community, who knew their money would be invested in cooperatives, and from the profits of the cooperatives themselves.

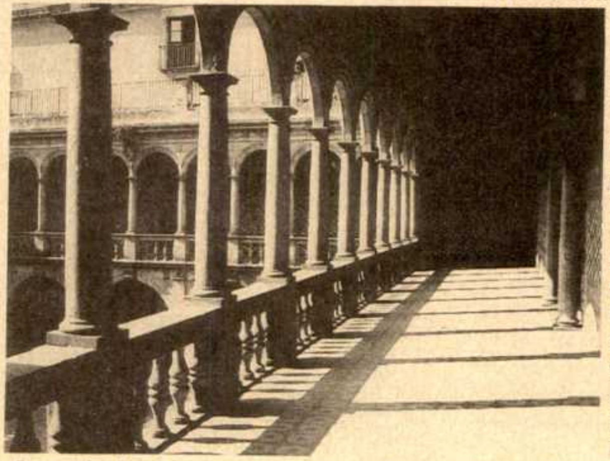
Each 'primary' cooperative is based on the simple principle that all who work there must be shareholders, and no one can be a shareholder who does not work there: the "indissoluble combination of worker and employer". The board of directors is elected and the highest-paid can earn no more than three times the salary of the lowest paid. The bank, is a 'secondary' cooperative, it provides services to all cooperatives, and has its management elected both by its own 1,000-plus workers and by the other cooperatives. The research institute, Ikerlan, set up in 1977, is another secondary cooperative.

The word 'cooperative' might conjure up an image of low-tech — of bearded idealists making handicrafts in pastoral surroundings — but that is not what Ikerlan is about (although the pastoral surroundings are there). Its aim is to develop the technology to make the cooperatives internationally competitive. Its laboratories are packed with computers, robots, experimental flexible manufacturing systems, and even a miniature 'robotized transportation vehicle'. It differs little from any small advanced technology institute from Tokyo to San Francisco.

Technological support is given to every cooperative (and other business) that solicits its services. Well over a hundred projects aimed at improving someone's product have been completed in its 12-year history. The research always involves members of the organization that requested it in order that training and technology transfer can be combined with product development — a philosophy that "was at first hard to get over to the government" which now provides for part of the running costs of Ikerlan. It is recognized that specific product development alone will not suffice to keep local industry competitive; half of Ikerlan's 70 scientists' time is now spent on generic research — basic work likely to later find specific application. In the computer sciences, for example, there are research programmes on geometric modelling, automatic programming of numerically-controlled machines and interactive computer-aided drawing. Just as elsewhere the aim is to provide a computer environment which helps an engineer to go from drawings for a specific product generated with the help of a computer to instructions for its computer-controlled creation. It is to the implementation of such research that small industries, producing say complex gears, look for their long-term health. ☐

Can there be a real 'Catalan science'

THE seventeenth century cloister shown here provides the entrance way to the Institute of Catalan Studies, maintained in rooms originally built as the house of convalescence to the fourteenth century hospital alongside. The institute is not as old as that, having been set up in 1907 on the crest of a revival in Catalan culture to aid "high scientific research" in the region.



Within a few years the institute had a structure rather like that of the old academies, with four sections: philology (mainly the study of the Catalan language), history and archaeology, science, and philosophy and social science. But its fortunes were thereafter to change with the political seasons, for any organization that might foster regional autonomy was suspect. Under the military dictatorship of Primo de Rivera (1923–30) it was suppressed, with the arrival of the Second Republic (1931–39) it flourished, and then when Franco came to power it had to continue its work of publishing scientific books in Catalan underground. Since 1976 things have returned to normality and the institute has returned to its original headquarters.

A large computer now hums away under the gothic arches in the basement helping to prepare a new dictionary of contemporary Catalan. For scientists, its principal importance is that all the local scientific societies are affiliated to it and rely upon it to help organize conferences and prepare publications. It has set up a Centre for Mathematical Research in the Autonomous University of Barcelona and is running joint projects; with the University of St Louis on the microseismicity of the Pyrenees, and with a Europe-wide group on the mapping of bryophytes.

The institute's president, Professor Enric Casassas i Simó, would like to see the institute expand further into a true interdisciplinary research centre. But there are others who see the institute's involvement in science as an anachronism. they argue that there is no place for "Catalan science" in the international age and particularly deride attempts to find Catalan equivalents for new scientific terms. ☐

Catalonia

Plenty of energy in Barcelona

THE really big news in Barcelona of course is not science but the Olympics. The choice of Barcelona as the site for the Olympic Games in 1992 sent Catalonia wild and will do much to heighten regional pride. It will also do something for science: a new biomedical institute is certain to be created for toxicological tests on the athletes and will live on after the games. And an Olympic village is going to be necessary and might afterwards provide much needed halls of residence for the universities (pessimists say the village will simply swallow up funds previously earmarked for the universities).

Like the Basque Country, Catalonia has a long history of independence and its own language and culture. But emotional appeals for independence are rarely heard, perhaps because Catalonia is recognizably prosperous and contains many of Spain's high-tech industries (as well as the Costa Brava). And it now has three excellent universities. In the Franco era the region was neglected. Despite the intellectual eminence of Barcelona, Spain's most cosmopolitan city with a population second only to Madrid, just four of CSIC's 78 institutes were located here: those for fisheries, humanities, geology and the Centre for Research and Development. New institutes are now springing up for microelectronics, materials science and economic analysis.

The region has its own commission for science and technology (CIRIT) and more than a few academics who would like to see Catalonia fully responsible for disbursing its own contribution to the central science budget. But until the unlikely day when Madrid relinquishes this power, CIRIT will have to make do with funds the Catalan government can raise to supplement central government allocations. All agree it cannot raise enough. Around 100 million pesetas goes each year in sending people abroad for training, and a further 200 million in research grants. Other funds are available for research in specific sectors, such as agriculture, but overall there is insufficient money for any large, specific project.

Of the CSIC institutes only the Centre for Research and Development is of any real size. As elsewhere, the institute has been through a painful process of reorganization. Just last year the five-storied building contained five institutes, with five directors, five libraries, and five administrators, all run separately. Now one institute, it brings together all the research groups in Barcelona working on pharmacology, organic chemistry and molecular biology. Changing the structure was clearly not easy: a bitter dispute over the election of laboratory heads has left one floor of the building with a crowd of researchers at one of its ends and two empty laboratories stripped of equipment, with one professor, formerly an institute chief, at the other. It is now up to the courts to decide whether the elections were held properly, and what belongs to who.

Sir Alexander Fleming is much fêted in Spain with streets or squares named after him in several cities. Above is a famous statue situated in a small square at the edge of the Barcelona red-light district. It is said that its inhabitants often leave flowers here in gratitude for his discovery of penicillin.

Overall director of the institute is Professor Joan Albaigés Riera. He sees the key advantage of the reorganization as creating groups of above critical size (at least 3 or 4 people in each, plus research students and technicians) and bringing related groups together to promote interaction. His speciality is environmental chemistry and entry into the EEC has greatly helped extend international connections and thus raise the level of research. The institute will be taking part in both the EUROTRAC and EUROMAR programmes to monitor air and sea pollution. □

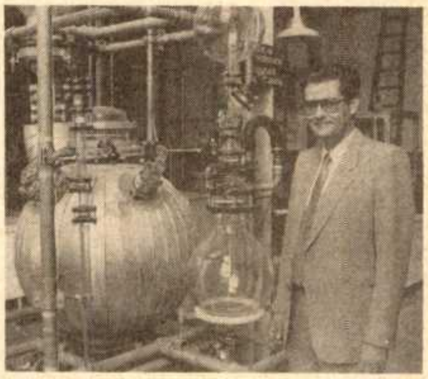
An education that pays for itself

IN an era of rapid change it is perhaps no surprise to find the occasional survivor from another age. Such is the Instituto Quimico de Sarria, set up in 1916, which can be summed as a "private Jesuit university with only one faculty — that for chemical engineering". State schools of engineering turned into polytechnics in the 1970s, and are now, if anything, more prestigious than the universities. Sarria continues as a single-subject university — but that does not mean it has lost prestige.

The institute, in an old suburb at the edge of Barcelona, shows its origins. At its centre are high-ceilinged laboratories, elegantly tiled in the fashion of the 1920s; across the corridor is a dimly-lit chapel entered by carved wooden doors.

But the institute is no fossil — courses are now well tailored to the needs of industry. A survey of its graduates finds 31 per cent of them working as directors in industry and another 42 per cent as technical executives. The institute's director, Dr Miguel Gassiot Matas, reckons there are two to three jobs available in industry for each graduate. The ability to guarantee their students a job seems a hallmark of all the Jesuit universities.

Part of the reason for the institute's success is its belief that it is not possible to teach well without doing research; a view that has only recently become fashionable in the state universities. Indeed the institute has to do research in order to survive; its costs are not met by the state and even with the high fees (around 300,000 pesetas, £1,500 a year) paid by its 600 students, it cannot make ends meet.



The director of the Institute Quimico de Sarria, Dr Miguel Gassiot Matas, in his teaching laboratory. The apparatus is not all just for demonstrating industrial processes; the students also help manufacture pharamaceutical caffeine which adds to the institute's income.

Almost a third of the institute's income stems from contract work, performed for industry and government. Much of the work is analysis of food, chemicals and pharmaceuticals but forensic science is also important. It was to the institute that the government turned during the outbreak of toxic oil syndrome, although it was a laboratory in London that discovered anilides in the oil a few days before them.

What of the future? The institute's director is well aware that each branch of science grows ever-more dependent on its neighbours; he looks to construct curricula connected with different fields to create "a small polytechnic school centred around chemical engineering". And with a look back to the political past he says, "an option to the state system is necessary; it is the most important guarantee of liberty". □

Portugal: political change

Dictatorship out with a whimper

JUST twelve years ago Portugal lacked a democratic government and was fighting wars in its African colonies of Mozambique, Guinea and Angola which had dragged on for a decade. A quarter of its male population of military age was in the armed forces. Now democracy is firmly established and of the colonies only Macao remains (until China takes it back).

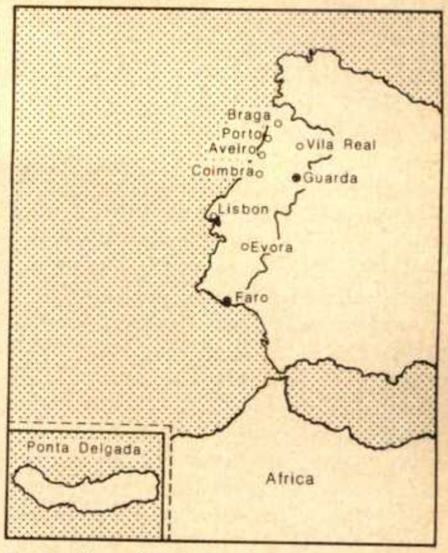
Although change has been rapid the recent relationship with the colonies still colours life more than in any other European nation. Almost a twelfth of the population has lived in Africa for when the military struggle was finally abandoned some 750,000 settlers came home. But as the Empire that stretched all the way east to Timor is becoming a memory, Portugal is turning to enthusiastically embrace Europe.

The series of political changes that led to the restoration of democracy, beginning with the 'revolution' in 1974, were triggered by weariness with the colonial wars. The dictator Antonio Salazar, who had ruled Portugal for 40 years, was already long out of office, having been paralysed by a stroke in 1968. But his successor, Marcelo Caetano, proved unable to modernize the country. Elections were held in 1969 but barely 20 per cent of the population were permitted to vote and opposition parties strictly supervised. The unsurprising result was a landslide victory for the ruling party.

Just how fragile the real power of the government had become was demonstrated on 25 April 1974. A group of army officers led a coup in which, with only a few shots and no casualties, over 40 years of dictatorial rule was brought to an end. Under the new rulers of the Armed Forces Movement power was placed in the hands of the Junta of National Salvation. At first, it took a moderate line but within a few months it lurched leftwards.

A Supreme Revolutionary Council was formed and began to implement radical policies whose effects are still apparent in Portuguese economic life. Large parts of the economy were nationalized, including banking, insurance, shipbuilding and transportation. In the south of the country, in the Alentejo, large private farms were seized by their workers and divided up. And the colonies were granted independence. But a massive turnout for the election of 1975 (to choose an Assembly to draw up a new constitution) gave support to the socialists, not the communists.

An unstable period followed, with attacks by communist sympathizers on the parliament, culminating in an attempt at a coup by left-wing military officers in 1975. Its failure proved the beginning of true democracy in Portugal. In the elections that followed the approval of the new constitution in 1976, the socialists formed the first freely-elected government. Since then the constitution has been revised giving Portugal an elected President, who appoints a Prime Minister, and an elected unicameral legislature, the Assembly of the Republic.



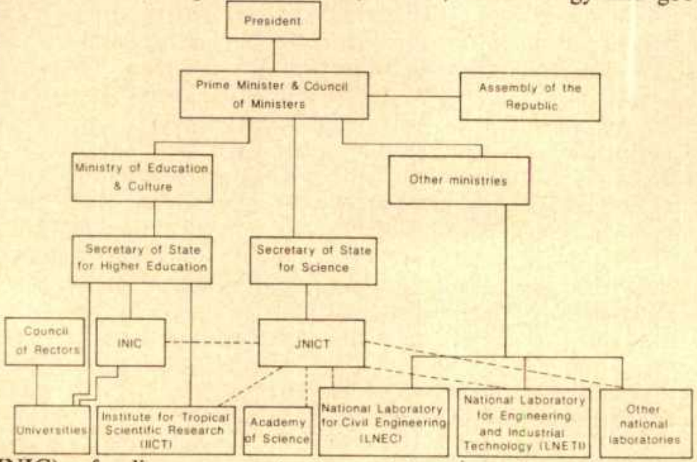
Portugal's twelve state universities: Azores (at Ponta Delgada), Algarve (at Faro), Aveiro, Coimbra, Évora, Lisboa (Lisbon), Nova de Lisboa, Técnica de Lisboa, Minho (at Braga), Porto, Beira Interior (at Guarda) and Trás-os-Montes e Alto Douro (at Vila Real). There are also a number of small private universities and a large Catholic University in Lisbon. The universities at Vila Real, Evora, and Faro are very new and are intended to act as poles of development in the inner regions of the country. Here conditions are backward and there is enormous scope for improvement in agriculture. Portugal has an illiteracy rate of 19 per cent and is still in desperate need of training schemes in basic technology. 90,000 students are now in higher education. A network of polytechnics is being set up to complement the universities.

To date no party has ever been able to gain an absolute majority and changes of government have been embarrassingly frequent. The Social Democrats have now overtaken the socialists as the largest group, although they are in turn easily outnumbered by the combined opposition (or by just the socialists and communists). There has been a general trend rightwards (the Supreme Revolutionary Council, which persisted as an advisory board was abolished in 1982) but the massive state holdings acquired earlier remain. Their status is highly controversial: they have performed poorly and the banks have proved unable to match the performance of smaller foreign entrants.

The lands seized in the Alentejo are gradually being returned to private hands as mistaken agricultural policies (including the replacement of the traditional cork trees by wheat) have failed. But the strength of the left prevents any rapid reversal of former policies. Key issues now are the need to raise the general educational standards (the illiteracy rate is the highest in Europe) and to improve agricultural practices in the poorer regions as barriers protecting Portugal from full exposure to EEC competition are gradually removed.

Pattern of research planning explained

THE Junta Nacional Investigaçáo Cientifica e Technologia (JNICT) is growing into the major body for the planning, promotion and coordination of Portugal's scientific and technological research. The overall control of JNICT goes to the Secretary of State for Science who operates under the jurisdiction of the Minister of Planning. Much of the research within the universities is carried out in centres that belong to the Instituto Nacional de Investigaçáo Cientifica (INIC), a funding agency operating under the Ministry of Education and Culture. Besides the twelve universities, one remarkable research institute, the Instituto de Investigaçáo Cientifica Tropical (IICT) is controlled by the education ministry. Other ministries control research institutes for agriculture, fisheries, health, meteorology and geophysics, and hydrographics. Best known are the Laborotório Nacional de Engenharia e Tecnologia Industrial (LNETI) and the Laborotório Nacional de Engenharia Civil (LNEC). □


President
Prime Minister & Council of Ministers
Assembly of the Republic
Ministry of Education & Culture
Other ministries
Secretary of State for Higher Education
Secretary of State for Science
Council of Rectors
INIC
JNICT
Universities
Institute for Tropical Scientific Research (IICT)
Academy of Science
National Laboratory for Civil Engineering (LNEC)
National Laboratory for Engineering and Industrial Technology (LNETI)
Other national laboratories

Science policy

A support system in evolution

Portugal is on the verge of taking scientific research seriously. Somewhere in its research organization exist bodies to monitor international trends and identify national priorities, evaluate fairly project proposals, coordinate different sectors, and encourage interaction between industry and academia — all the things a modern nation needs. And it has a supply of well-trained researchers. The problem is that the various parts are not all the right size.

Portuguese scientists are themselves well aware of the problems — even enthusiastic about tackling them. Change has begun. In the scenario now developing two organizations will certainly be important: the National Board for Scientific and Technological Research (JNICT) and the National Institute for Scientific Research (INIC). The former's star is on the rise while the latter, even by the admission of those who run it, could profitably be reformed in a more sophisticated organization.

JNICT

As the Minister of Education, Professor João de Deus Pinheiro, put it, "JNICT was the turning point". When it was set up in 1967 it was the first organization to evaluate projects using independent referees, including foreigners, and to follow up and assess how well projects were running. Now JNICT appears as the one national body planning, coordinating and funding scientific and technological research in all sectors of activity. But it is still not the real scientific power in the land it could be, despite its 20 year history.

Two things have held JNICT back: frequent changes in its position in the administration (it is now in the Ministry of Planning, under a Secretary of State for Scientific Research, but this is its sixth resting place) and, more important, the failure of the government to give it sufficient funds. The total science budget now runs close to 13,000 million escudos (65 per cent of which is government expenditure, most of the remainder coming from industry) but last year JNICT received only 200 million escudos (1.5 per cent of the total).

The future looks brighter. This year the budget rose to 1,000 million escudos and next year it should go 2,500 million escudos (19.2 per cent of the total), for the first time putting it ahead both of INIC and the largest public research institute, the National Laboratory of Engineering and Industrial Technology (LNETI).

JNICT's grant can flow anywhere within the research and development system — to universities and their research centres, to state laboratories, to industry, and to non-profit research organizations. It is thus well placed to ensure that national priorities are met and gaps filled. In the words of the JNICT president, Professor José Mariano Gago, "the strongest indication of a trend to increase overall coordination is the amount of money being put here." A more powerful JNICT with what Professor Pinheiro called a dual policy of backing "both the best researchers and the areas that give an advantage to Portugal" could be just what the country needs. With its bigger budget its plans will grow. Next year the largest slice will go on contract research, concentrated in four priority areas: material science, marine science, biotechnology and electronics. Part of the remainder will go to encouraging industrial applications of research but a sizeable portion will be reserved for "encouraging excellence": ten to fifteen groups of real international standing will be supported whatever their field of study might be.

One problem that JNICT is trying to cure is that of "small-scale groups who can never achieve good results in time". Attempts are being made to create new networks of researchers. Membership of CERN is providing some opportunities for at first the annual contribution will be spent in Portugal building up resources. A Lisbon group specializing in microcomputers and optical communication will link up to a Coinbra group with expertise in nuclear detectors and instrumentation to form a new 'institute'. The share out of funds is entirely decided by an international committee. A similar scheme will create a new institute of biotechnology.

INIC

For the universities the Ministry of Education's own grant body, INIC, is of prime importance in supporting research. That support does not, however, go directly to the universities but mostly to those in INIC's own 126 research centres within the universities. The 3,500 researchers in the centres are almost all university professors so their salaries come from a separate source but they have fewer teaching and administrative duties that staff outside the centres. That leaves both INIC and the universities in a curious position; INIC has a large number of quite small research centres whose staff's promotion is decided by the university, while the universities have no real control over the research activities of their own staff. Worse still is that research centres covering the full range of disciplines have to be maintained and that means INIC is under obligation to continue funding them all by block grants. Projects assessment is included in the complicated formula used to calculate each centre's annual budget but there is no strong sense of competition for research funds and, in any case, INIC's referees are not noted for their critical powers.

Originally, this system made sense in that central distribution helped overcome the universities' own problems in sharing out their budgets. But as the universities have matured and a Law of University Autonomy looms on the horizon there seems little point in the present system. INIC officials are themselves proposing change. Many of the centres, they say, could be handed over to the universities and instead some twenty or so larger interdisciplinary institutes be created. The new organizations springing up to serve high-energy physics and biotechnology could provide a model. The problem is that changes of this kind have been suggested for some years, without much happening. The President of INIC, Alberto José Correia Ralha, a very patient man, says "sometimes even I think I am going to get angry". ☐

Portugal's Minister of Education, Professor João de Deus Pinheiro has been in his job two years "an almost unimaginably long time" given the frequency with which governments change. Total expenditure on research he sees as low although it is probably double the official figure of 0.4 per cent of GNP because "salaries are not included". But change is going to involve more than just increasing budgets. "For the fifty years of authoritarian government a good student was one who could repeat the text book or the teacher's words. Creativity and criticism were words absent from our education system" he says. A massive review of the education system is now under way in which priority will go to improving science teaching in the schools through teacher retraining. To improve the research system help is being sought from the World Bank for institutes in biotechnology and agriculture, water resources, and materials. An institute for marine science is also planned. Improving links with industry is a priority. Interministerial coordination is still poor and relations with the Ministry of Industry have "not always been good".

Young Portuguese, the minister says, might learn something from thinking afresh about history. Rather than viewing the fifteenth century voyages as an outcome of the desire to spread Christianity they can be seen as stemming from wise technological policy. Portugal's breakthroughs in ship design and organization created the equivalent of "today's space rockets" and made exploration possible. ☐

Tropical research

Everything under the Sun

The only way to describe Lisbon's remarkable Instituto de Investigação Científica Tropical is a 'horizontal institute'. Rather than specializing in one field, it covers almost everything provided it can be related to the tropics. Geology, archaeology, zoology, ethnology, geodesy, veterinary medicine, history, ancient maps, soil science and anthropology are studied among other things, and there is a botanical garden.

There is a historical explanation for the institute's structure. It began life just over a hundred years ago as a Commission for Cartography to map the new colonies and, as the colonies grew, so it expanded to offer them a complete range of scientific services. And if it had not been for the British, the colonies would probably have been a lot bigger. Maps produced early this century show everything between Angola and Mozambique painted the 'rose colour' of Portuguese territory. But the British sternly warned that their north-south expansion was not going to be halted by a Portuguese east-west connection. The resulting crisis and threats that gunboats were to be dispatched brought down the government of the day.

After the colonies were given up, twelve years ago, there was a series of upheavals (along with changes of name from 'Junta' to 'Laboratorio' to 'Instituto') leaving the institute in the care of the Ministry of Education with the aim of aiding scientific research in the tropics (in 40 countries, not just the former colonies) and training specialized personnel. The government budget is 800 million escudos a year but further help comes from international organizations; the United Nations, UNESCO and the African Development Bank among them. The different specialities are not housed under one roof but in independent centres and each is connected to relevant university departments: the directors are also university professors.

For the international community perhaps the most important centre is that for the Investigação das Ferrugens do Cafeeiro (CIFC), a world centre for studying coffee rusts — fungal diseases. Coffee may be only a drink but it is of tremendous economic importance: second only to petroleum in the value of its world trade.

Rusts are important because of what they could do to the industry. The major commercial species of coffee is *Coffea arabica*. But *arabica* is highly susceptible to coffee leaf rust, which quickly destroys the plant. In the first major outbreak, towards the end of the nineteenth century, Ceylon's coffee industry was utterly destroyed. Little natural resistance has been found among varieties of *arabica*. And to make matters worse, it appears that all the coffee plantations in Central and South America, where 65 per cent of the world supply is grown, were established from the progeny of a single plant, taken from Java to the botanical garden in Amsterdam in 1706.

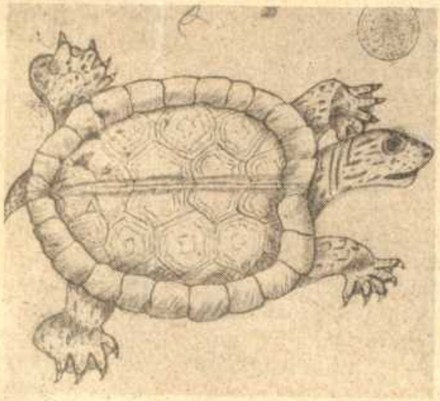
Angola, Portugal's biggest African colony, was the world's third-greatest coffee producer. But the coffee rust research centre in 1955 owes less to this link than to US worries about what an outbreak of coffee rust could do to the economic stability of Latin American countries dependent on coffee crops. In 1952 two US scientists set out on a world-wide tour to look at the disease and, after encountering a Portuguese researcher, decided Lisbon would be the ideal place for a centre — no coffee was grown in Portugal and rusts from around the world could be studied safely there.

Coffee is big business and a fungus, coffee rust, a big threat to it. Here just a few of the tens of thousands of coffee plants that pass before the Centro de Investigação das Ferrugens do Cafeeiro director C.J. Rodrigues Jr for testing each year.

In those days, the fungus had not reached the Americas. But in 1970 the first outbreak occurred in Brazil, and within a few years the rust had spread all over the country. Spores are probably carried by the wind for they have been found 2,000 metres up in the air. By 1983 rust had crossed the Andes into Columbia and also appeared in Central America.

In the intervening years the centre had accumulated some 33 different races of the fungus and a vast collection of coffee varieties that had enabled them to classify plants according to their resistance spectrum. With this standardization done they could begin work on the genetics of resistance. And they had discovered in another of their colonies, Portuguese Timor, a hybrid between *C. arabica* and *C. canephora* that seemed totally resistant to almost all fungal races. Furthermore, like *arabica* it was tetraploid and self-fertile. Luckily the resistance genes were dominant. A programme of crossing and backcrossing and resistance testing then began to introduce them into an *arabica* background.

Altogether nine genes for rust resistance have now been identified. The centre is cooperating with researchers all around the world to help develop resistant strains, with the 'Timor hybrid' still at the centre. Breeders send tens of thousands of coffee plants by air every year for resistance testing in the centre's greenhouses — for many growing regions the choice may soon become resistant strains or crippling bills for fungicides. Plans are now afoot to use cryogenic techniques to preserve spores and to develop tissue culture methods so that known lines of coffee plants can be propagated quickly. □



Given its long history and colonial links it is not surprising that the Institute for Tropical Science has amassed some treasures. Above is an illustration showing a turtle from one of the earliest works on the natural history of Brazil, the *História dos Animals e Arvores do Maranhão*, by Father Cristóvão de Lisboa who was resident in Maranhão in 1627. The monograph is on display at the Archives of Overseas History which now houses the records of the old Overseas Territories Office as well as other collections.

Another treasure house, too little known outside Portugal, is in the Centre for Ethnology. For years it has put its collections on public view only rarely but now it is rigorously pursuing an open policy. An astonishing collection of African sculpture is now on view (all, alas, the product of cultures already destroyed) and in the basement store-room glass fronts are being added to crowded storage shelves so that the public will be able to see everything the museum has. □

Portugal's university research

Adding a little more flexibility

Research funds from both INIC and JNICT are available to university researchers. But according to a group of scientists drawn from the New University of Lisbon's Faculty of Science and Engineering, the total does not necessarily add up to much: "Just a little from each, only occasionally a jackpot". In fact, in a good year there are around 120 million escudos (£600,000) to support the research of the 70 staff, including money from industrial sources.

This university is now trying, with a lead provided by Professor Leopoldo Guimarães, to create an alternative in UNINOVA, an institute bringing together the universities and industry. Industry has not been totally neglected in the past. The faculty's Department of Environmental Sciences real research is paid for by contracts — a bimass-from-wine production effluents project paid for by a German company, and environmental assessements projects from the Ministry of Industry and the local municipality. UNINOVA — the Institute for the Development of New Technology — is intended to be on a different scale. Some 16 companies, including ICI, Sperry and Digital, are taking shares in the institute. The idea is that the involvement of Portugal's biggest (and often foreign) companies in electronics, computers and biotechnology will power research in those areas as well as providing funds. The institute is not intended to perform research directly for the companies involved.

A simialar organization has already proved a success at the Technical University of Lisbon, and the vice-rector there, Professor Anotónio Simões Lopes, is particularly proud of it. It also has the Institute for Systems and Computer Engineering set up jointly by the university, the University of Porto and the three telecommunication operators. Just six years old it already provides for 300 researchers.

Such institutes will add new flexibility to the university system. It has not proved easy in the past to add numbers where they are most needed, partly because PhD students are virtually guaranteed jobs in the same university so there are limits on how many of them can be taken in.

Despite these difficulties it would be wholly wrong to think of university staff as demoralized. Up to the late 1970s many people were sent abroad for doctoral training, particularly to Britain, and have provided Portugal with a distinctive core of enthusiastic young (under 40) researchers. Every well-established university has a group of international standing. Recent successes (in international collaborations) include the isolation of LAV II virus (with Luc Montagnier of the Institute Pasteur) and the controversial proof of Poincare's conjecture (with Colin Rourke of the University of Warwick). And Coinbra's nuclear detector group has even attracted a doctoral student from Japan. ☐

Industry

No shift from ivory towers

Professor Clemente Pedro Nunes, Director General of Higher Education, is new to his job and bursting with energy. He is a chemical engineer, with a doctorate from the University of Birmingham, and plenty of experience of working in industry. The great problem he wants to tackle is the failure of the universities and industry to interact. One statistic sums it up: there are now over 2,000 holders of doctoral degrees in Portugal but less than 10 are working in industry. And there are no signs of a trend to move into the industrial sectors, except in computer software.

Much of the blame must lie with industry for failing to invest in research and development. "Most of all what is needed", says Professor Nunes, "is managers in industry who can discuss science and technology; even in big companies there are very very few." But the universities are also at fault as the progression from a masters degree through a doctoral degree to lifetime tenure is almost assured. There is no pressure to leave and "naturally every professor wants new blood in the department and knows every trick in the book to force extra vacancies. The bill, of course, comes to us (the Ministry of Education). And university salaries are not much different from those in industry."

Professor Nunes wants more scientists in industry.

Fierce words, perhaps, but how can industry be persuaded that research and development is in its best interests? Tax incentives and perhaps "a sabbatical year in industry paid by the state" suggests Professor Nunes. Part of the problem is that "industry is undercapitalized. Nationalization in the mid 1970s broke the back of many structures and they have not yet recovered". Encouraging signs are coming from the joint university–industry research institutes set up within universities. Polytechnic universities also hold out promise for raising industry and agriculture's level of sophistication. There are now 16 in existence, but not all are complete and last year they could take in only 1,000 students. But this year is to be the "big year for their launch". ☐

Oiling the wheels

"It's difficult to imagine what Portuguese intellectual life would be without the Gulbenkian Foundation" said Ernesto Veiga de Oliveira, ex-director of Lisbon's Museum of Ethnology; a comment that well summed up what many other researchers had said about the foundation.

The foundation gives away more than seven thousand million escudos each year and even though science is the least of its benificiaries (receiving 12.6 per cent of the total; after education, arts and social welfare) its impact is clearly felt in the research community. But just why is the Gulbenkian Foundation based in Portugal?

It is a complicated tale. Calouste Gulbenkian was born in Istanbul but took British nationality and was a graduate of Imperial College, London. There he gained a knowledge of mining engineering that helped him realize the potential importance of Middle East oil. By the time he died in 1955 he owned 5 per cent of Iraq's oil reserves. In between he made his home in Paris but, in 1942, after the German invasion, he moved to Portugal, which remained neutral (although favouring the Allies) during the Second World War. He hired a hotel (now the Sheraton), brought in a couple of Rolls-Royces and, finding Portugal to his taste, continued to live there after the war. Along the way he collected the paintings, furniture, ceramics and silver that now form the collections of Lisbon's biggest museums.

The foundation played a key role in Portugal's scientific development. Starting in the mid-1960s, when the universities were at a low ebb, hundreds of scientists were given grants to study abroad. Many now occupy important positions in the universities and government. An institute of science was also established but as the universities developed, the need for it has lessened and it is gradually being closed down. Fields of strategic importance (informatics is now the centre of attention), are helped with special equipment grants. Translations of technical works are made so that inexpensive editions can be published alongside books from Portuguese authors. And links with other Portuguese-speaking countries are kept up through special cooperation agreements. ☐

Superfluid ^{3}He in rotation

V. P. Mineev*, M. M. Salomaa† & O. V. Lounasmaa†

* L. D. Landau Institute for Theoretical Physics, USSR Academy of Sciences, 117334, Moscow, USSR
† Low Temperature Laboratory, Helsinki University of Technology, SF-02150 Espoo, Finland

Experiments on vortices in rotating superfluid ^{3}He have revealed phenomena which have not been observed in any other quantum liquid and have implications in fields beyond low-temperature physics. In ^{3}He-A, vorticity may be supported by a continuous winding of the order parameter which gives rise to the observed continuous coreless vortices with two flow quanta. Exotic vortices with half-integer circulation quanta may also exist in this phase. In ^{3}He-B, an abrupt change in the vortex-core structure has been explained theoretically as a topological phase transition involving half-quantum vortices. NMR experiments on rotating ^{3}He-B have revealed that the cores of vortices observed possess spontaneously broken parity, manifested by their ferromagnetic superfluid cores.

QUANTIZED vortices in rotating superfluid ^{3}He display novel structures that make them interesting objects in condensed-matter physics. This can be attributed to their anisotropic nature and rich internal structure of ordering: unlike superfluid ^{4}He (He II), the superfluid phases of liquid ^{3}He are not only superfluids but also liquid crystals, magnets, and spin liquid crystals; a few, or all of these different aspects of ordering may contribute to the properties of inhomogeneous states, such as quantized vortices.

Superfluidity and superconductivity are observed in liquids whose macroscopic behaviour is governed by quantum statistics; the latter becomes important at low temperatures when the thermal de Broglie wavelength of the particles in the medium exceeds the interatomic spacing. Superfluids can flow without friction through narrow gaps and tubes. Only two liquids, ^{4}He and ^{3}He, are known to exhibit this property. Superconductivity, the loss of electrical resistance, is observed in many metals below a critical temperature T_c that varies from one substance to another. In superconductors, superfluidity is manifested by the conduction electrons.

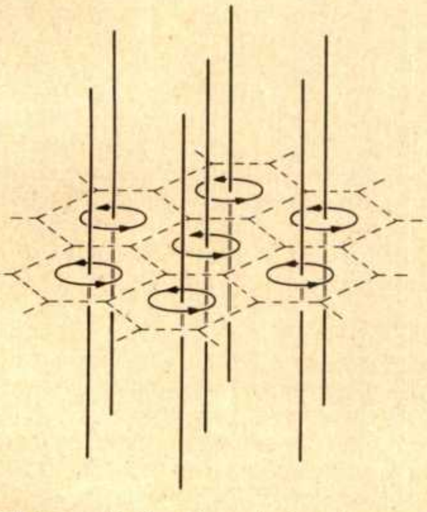


Fig. 1 The hexagonal lattice of quantized vortices, as seen by an observer co-rotating with the superfluid. The rotation axis is parallel to the array of vertical lines depicting vortex cores at the centre of each primitive lattice cell; lines with arrows illustrate streamlines of superflow. Dashed lines indicate the hexagonal vortex-lattice cell boundaries, along which the superfluid is at rest in relation to the rotating container. On average, the lattice of quantized vortices mimics solid-body rotation with a parabolic meniscus. Rotating superfluid ^{4}He has been studied since the 1940s[1], first by Andronikashvili[2] extensive measurements have been conducted recently by Glaberson, by Packard, by Reppy, by Rudnick, and by Tsakadze and their co-workers.

Phenomenologically, a superfluid can be considered a liquid with two components: the normal fraction has density ρ_n and the superfluid fraction has density ρ_s. The superfluid behaves like a single macroscopic particle with a quantum-mechanical wavefunction $\Psi(\mathbf{r})$, usually called the order parameter. Physical properties of particular interest in superfluids are those in which the quantum character of the liquid manifests itself on a macroscopic scale. The most interesting of these are phenomena that appear in superfluid ^{4}He and ^{3}He under rotation. Due to their quantum nature, unlike classical liquids, superfluids cannot rotate as a whole. Instead, circulation perforates them in quanta of individual vortex lines[1,2] (Fig. 1).

In 1972, Osheroff, Richardson, and Lee found the superfluid A and B phases of ^{3}He below 3 mK; the P/T (pressure/temperature) diagram is illustrated in Fig. 2. This discovery opened a new and exciting era in ultralow-temperature physics. In 1981, vortices in rotating superfluid ^{3}He were first investigated experimentally at the Helsinki University of Technology[3].

During the past five years, intensive studies with the nuclear magnetic resonance (NMR) technique, persistent-current measurements, ion-mobility investigations, and hydrodynamic experiments have all revealed several new, unprecedented and often unexpected phenomena associated with vorticity in anisotropic superfluids. Space permits only a brief discussion of persistent-current experiments and a more detailed description of the pioneering NMR results; the latter have provided most of the fundamental information on vortices in superfluid ^{3}He. For example, in ^{3}He-A, continuous distribution of vorticity has been observed, as in a classical liquid—accompanied, however, by a liquid-crystal-like texture. In ^{3}He-B, new superfluid phases nucleate inside the cores of the quantized vortex lines—as revealed by the spontaneous vortex-core magnetization and the first-order vortex-core phase transition[3] (see Fig. 2).

Vortices in classical superfluids

In ^{4}He, which has a nuclear spin $S=0$ and which obeys Bose–Einstein statistics, superfluidity is understood as the Bose condensation of individual atoms into the zero-momentum state[4,5]. The Bose condensate may be described by a macroscopic wavefunction $\Psi(\mathbf{r})=C(\mathbf{r})e^{i\Phi(\mathbf{r})}$, whose amplitude squared, $C^2(\mathbf{r})$, is the density, $\rho_s(\mathbf{r})$, of the superfluid component. The

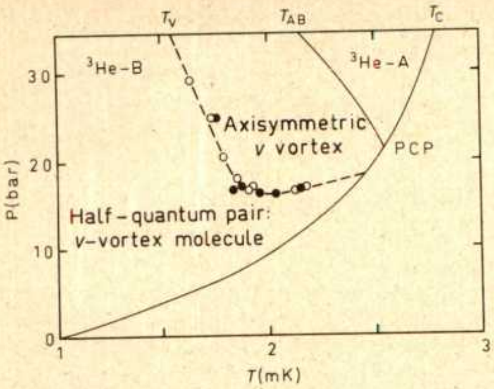


Fig. 2 Phase diagram of liquid ^{3}He. On rotating ^{3}He-B in an open-volume NMR cell, a transition is observed[3] on crossing the dashed line T_v in the P,T-plane[15]. The intersection of the T_{AB} and T_c curves is the polycritical point, PCP. ○, NMR experiments at $H = 284$ G; ●, $H = 568$ G. This vortex-core transition is a topological change between two inequivalent ways to accommodate vorticity[48,49]; see Fig. 8.

gradient of the phase, $\nabla\Phi(\mathbf{r})$, gives the superflow velocity:

$$\mathbf{v}_s = \frac{\hbar}{m_4}\nabla\Phi(\mathbf{r}) \tag{1}$$

Being the gradient of a scalar, superflow is potential, curl-free: $\nabla \times \mathbf{v}_s = 0$. Experiments show that this is true everywhere, except along special vortex lines at which $\nabla\Phi$ is singular; the flow velocity diverges as $v_s \propto 1/r$ where r is the distance from the vortex core. The integral $\oint \mathbf{v}_s\, d\mathbf{r}$ along any closed path around a vortex is proportional to the change in phase $\Delta\Phi$, which—to keep the wavefunction single-valued—must equal an integral multiple of 2π. Hence the circulation of superflow is quantized in He II:

$$\oint \mathbf{v}_s\, d\mathbf{r} = m\frac{h}{m_4} \tag{2}$$

in quanta of $h/m_4 = 0.997 \times 10^{-3}$ cm^2 s^{-1}; m_4 is the mass of a ^{4}He atom and m is an integer. Equation (2) resembles the Bohr–Sommerfeld quantization condition: the angular momentum per particle is quantized in units of $\hbar$.

Viscous fluids in a container rotating at the angular velocity $\boldsymbol{\Omega}$ behave like a solid body, that is $\mathbf{v} = \boldsymbol{\Omega} \times \mathbf{r}$ and $\nabla \times \mathbf{v} = 2\boldsymbol{\Omega}$. This is not possible for a superfluid, but because rotational motion is energetically more favourable than the stationary nonrotating situation, a vortex lattice (Fig. 1), is created.

Each vortex normally carries only one quantum of circulation, because the energy of two vortex lines is smaller than the energy of a single one with two circulation quanta. Repelling each other, vortices form a hexagonal two-dimensional lattice. On average, such an arrangement imitates solid-body rotation. Consequently, the total number of singly quantized vortex lines in a vessel of cross-sectional area A is

$$N = \frac{2\Omega A m_4}{h} \tag{3}$$

and this corresponds to $N/A \simeq 2{,}000$ vortices cm^{-2} at $\Omega = 1$ rad s^{-1}.

Vortex lines in superfluid ^{4}He have no particular internal structure: the superfluid density tends to zero on the vortex axis, where the fluid is normal. Hence the kinetic-energy density of the rotating superfluid component, $\frac{1}{2}\rho_s v_s^2$, remains finite everywhere, in spite of the $1/r$-singularity of $\mathbf{v}_s$ on the vortex axis. The characteristic scale for spatial changes in the order parameter, determining the radius of the vortex core, is the so-called coherence length, $\xi = 0.1$–0.2 nm in ^{4}He; this is of the same order as the interatomic distance.

Superfluid phases of ^{3}He

Like electrons, ^{3}He atoms having a half-integer (nuclear) spin $S = \frac{1}{2}$ are governed by Fermi-Dirac statistics. Strong mutual interactions renormalize the ^{3}He atoms into quasiparticles but do not change the nature of the quantum statistics. According to the BCS theory of Bardeen, Cooper and Schrieffer, Fermi systems undergo a transition into a macroscopically coherent new superfluid ground state in the presence of an arbitrarily weak attraction between the quasiparticles. The long-range order may be viewed as being due to the formation of mutually overlapping giant molecules, called Cooper pairs.

In superconductors[6], as a consequence of the isotropy of interactions between the conduction electrons, these pairs have angular momentum $L = 0$ (s-wave pairing) and total spin $S = 0$ (singlet). In superfluid ^{3}He, the Cooper-pair states[7] have $L = 1$ (p-wave pairing) and $S = 1$ (triplet). The order parameter $\Psi(\mathbf{r})$ for a p-wave paired superfluid can be written as a linear combination of spherical harmonics χ_{S_μ} and $Y_{L\nu}$, respectively, for the states with different projections $S_\zeta = \pm 1, 0$ ($\chi_{1\mu}$ with $\mu = \pm 1, 0$) and $L_z = \pm 1, 0$ ($Y_{1\nu}$ with $\nu = \pm 1, 0$) on directions of the quantization axes, $\hat{\zeta}$ and $\hat{z}$:

$$\Psi(\mathbf{r}) = \sum_{\mu\nu} a_{\mu\nu}\chi_{1\mu}Y_{1\nu} \tag{4}$$

In ^{3}He-A, all pairs are in the state with $L_z = 1$ and $S_\zeta = 0$, that is, there is only one component, $a_{0+} \neq 0$. In ^{3}He-B, with $L + S = J = 0$, states with $L_z = \pm 1, 0$ and $S_\zeta = \pm 1, 0$ are equally populated, and there are three equal components of the order parameter: $a_{+-} = a_{00} = a_{-+}$.

In both the A and B phases[8,9], the average value of the spin, namely the spontaneous Cooper-pair magnetization, is zero. In ^{3}He-B, the average of the angular momentum is also zero; hence this liquid is completely isotropic in the absence of inhomogeneities. In contrast, all Cooper pairs in ^{3}He-A share the same direction of $\mathbf{L}$, which determines the orbital anisotropy axis $\mathbf{l}(\|\hat{\mathbf{z}})$. Therefore, many properties of ^{3}He-A, for example viscosity, are different in directions parallel or perpendicular to $\mathbf{l}$. Hence ^{3}He-A is an anisotropic superfluid liquid crystal. In addition to the orbital anisotropy, there is a magnetic anisotropy in the A phase, described by the vector $\mathbf{d}$ which defines the spin quantization axis $\hat{\zeta}$; the Cooper-pair spin $\mathbf{S}$ is perpendicular to $\mathbf{d}$. Because of this anisotropy, the paramagnetic susceptibility of the A phase is a uniaxial tensor.

Superfluidity in ^{3}He-A is not necessarily associated with curl-free flow because $\nabla \times \mathbf{v}_s$ may be different from zero. In fact, the A phase can support nonsingular vorticity in the presence of an inhomogeneous distribution of the anisotropy axis $\mathbf{l}$. Previously, curl-free flow was considered an essential property inherent to every superfluid. Thus it has become necessary to revise many theoretical concepts of quantum liquids, especially their behaviour under rotation.

Rotating cryostat

All NMR measurements—and most of the other types of experiments—on superfluid ^{3}He in the rotating equilibrium state have been made in the ROTA cryostat at Helsinki[10,11]. This apparatus incorporates a copper nuclear stage, a dilution refrigerator, an activated charcoal cryoadsorption pump, radial and axial air bearings, and computerized measurement electronics. The apparatus can be rotated at angular velocities Ω up to 3.0 rad s^{-1}. The copper nuclear stage, weighing 4.0 kg, is first precooled to 20 mK, with the 8-T magnet on, by means of a dilution refrigerator. The nuclear stage is then thermally isolated by operating a superconducting heat switch and demagnetized in 10 h to 150 mT, whereby it cools to 0.4 mK. Next, the charcoal adsorption pump is switched on and the external pumping lines are disconnected, allowing the cryostat to be rotated.

The ^{3}He sample cell is on top of the nuclear stage and thermally connected to it. For NMR measurements the cell was a cylinder, 5 mm in diameter and 30 mm long. Its geometry is thus clearly defined, and the NMR response of stationary ^{3}He in this configuration is well known. The measurements were made at three fixed frequencies; 1,845, 920 and 460 kHz. The magnetic field was usually oriented along the rotation axis but could be tilted sideways when required. The field was swept linearly in 2–3 min over the resonance at 56.9, 28.4 or 14.2 mT. For persistent-flow measurements, the experimental cell was a 50-mm-diameter torus, 6-mm thick; the flow path was filled with 20-μm plastic powder to enhance v_c, the critical velocity for superflow.

Persistent currents in ^{3}He

One of the characteristic properties of superfluids is their persistent-flow capability: in He II the flow of the superfluid component around a closed path persists as long as the liquid is kept under its transition temperature, $T_\lambda = 2.17$ K. Experiments for observing persistent currents in the superfluid phases of ^{3}He have been carried out by groups at Cornell University[12,13] and in Helsinki[14,15]. In both laboratories, an a.c. gyroscopic technique was used. Figure 3 illustrates the gyroscope employed in the Cornell experiments. The principle of the measurement is: when the ring is driven about the y-axis at the resonant frequency of the ω-mode, a sinusoidal response around the x-axis results, provided that there is an angular momentum in the ring which is caused by superfluid flow.

In the gyroscopic experiments, the sample was first cooled well below the superfluid transition temperature at rest. The cryostat was then rotated around the z axis at an angular velocity Ω and stopped, and the vibrational amplitude in the ω-direction was recorded. An angular momentum was detected in the torus only if the speed of rotation had exceeded a critical value Ω_c. In such conditions, an irreversible creation of superflow occurs which persists even after the cryostat has been brought to rest. In ^{3}He-B, no decay of superflow was observed in Helsinki[14] during 48 h, implying an effective viscosity at least 12 orders of magnitude less than that in the normal Fermi liquid at the same temperature: thus ^{3}He-B shares properties with 'classical' superfluids. In the Helsinki experiments no observable current was found to persist in ^{3}He-A. However, the Cornell group[13], exploiting a method with a higher quality factor near T_c, also detected persistent currents in ^{3}He-A. A discontinuous change in the critical flow velocity, attributed to a vortex transition in ^{3}He-B was seen in the Helsinki experiments[15].

Vortices in ^{3}He-A

Before any experiments had been conducted on rotating ^{3}He-A, it had been recognized on theoretical grounds that this phase should possess curious superflow properties[16–18]. In particular, the possible existence of continuous vortices—which were subsequently found in Helsinki—had first been predicted by Chechetkin[19] and also discussed by Anderson and Toulouse[20]. In addition, half-quantum vortices had been suggested by Volovik and Mineev[21] and subsequently by Cross and Brinkman[22] on the basis of the Cooper-pair structure in ^{3}He-A.

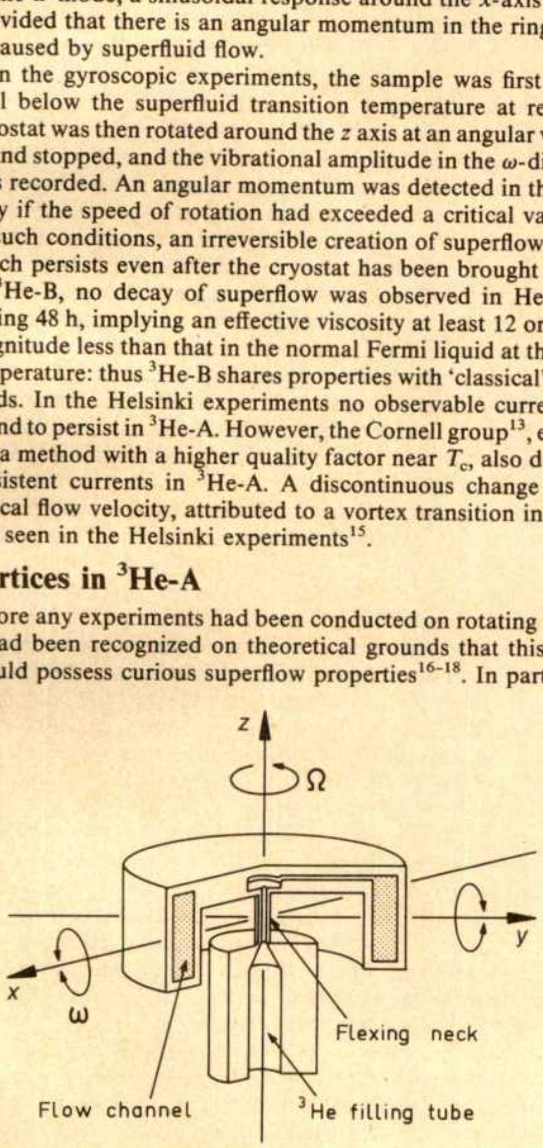


Fig. 3 Schematic diagram of the a.c. gyroscope employed in the Cornell experiments[12]. To immobilize the normal fluid, the toroidal flow channel is filled with fibrous material; the resulting average open volumes have linear dimensions of the order of 0.1 mm. For operating details, see text.

Superfluidity of ^{3}He-A

The p-wave pairing state of ^{3}He-A was discussed above: each Cooper pair in this liquid is under rotational motion with the projection $L_z = 1$ of the angular momentum on the direction of the quantization axis $\mathbf{l}$. This internal motion is mixed with the macroscopic Cooper-pair centre-of-mass motion. Because of this, the nature of superflow in ^{3}He-A is very peculiar: rotational motion can follow from the presence of spatially inhomogeneous distributions of the $\mathbf{l}$ vector[17]. These textures unbalance the internal Cooper-pair motion, producing a nonpotential net superflow without the formation of singularities.

The spatial structure of the order parameter is determined by competition among several forces. The elastic stiffness of the $\mathbf{l}$ and $\mathbf{d}$ fields tends to make them uniform; this gives ^{3}He its liquid-crystal-like behaviour. The dipolar interaction tries to lock $\mathbf{l}$ and $\mathbf{d}$ parallel or antiparallel to each other. At cell boundaries, $\mathbf{l}$ is perpendicular to the walls[16]. In the presence of superfluid flow, $\mathbf{l}$ tends to orient parallel to $\mathbf{v}_s$. An external magnetic field tries to lock $\mathbf{d}$ into the plane perpendicular to $\mathbf{H}$ ($\mathbf{d} \perp \mathbf{S} \parallel \mathbf{H}$).

The actual texture of the liquid depends on the interplay between these effects, with the configuration of lowest free energy prevailing. As the textures of $\mathbf{l}$ can create rotational superflow in ^{3}He-A, the circulation of superflow velocity in the A phase need not be quantized. At first sight, the possibility of nonpotential superflow seems to admit purely solid-body rotational motion. However, the energetics of the $\mathbf{l}$ texture require that even in ^{3}He-A under rotation, a periodicity in the order-parameter texture appears.

Vortex structures

Every elementary vortex-lattice cell contains a vortex line with m quanta of circulation. The value of the circulation quantum in superfluid ^{3}He, $h/2m_3 = 0.662 \times 10^{-3}$ cm^2 s^{-1}, differs from that in superfluid ^{4}He [see equation (2)] not only due to the different isotopic mass, but also owing to the factor 2, caused by Cooper pairing.

The possibility of rotational superflow means that a singular vortex line can be transformed into another one with continuous coreless distribution of vorticity. However, due to topological constraints[23], the elimination of the singularity in $\mathbf{v}_s$ on vortex lines with odd number of circulation quanta is accompanied by the appearance of a singularity in the $\mathbf{l}$ texture ($\nabla \mathbf{l} \to \infty$ on the axis). However, vortices with an even m can be transformed into vortices with a continuous distribution of vorticity, while maintaining a nonsingular distribution of the $\mathbf{l}$ texture. Hence, vortices with odd m are always singular, but the singularity in vortices with even m can be continuously dissolved.

The singular singly quantized vortex has two different cores, the hard core with radius of the order of the coherence length, $\xi \simeq 10$ nm, inside which the order parameter deviates from its bulk ^{3}He-A value, and the soft core, with a radius of the order of $\xi_D \simeq 6$ μm, the dipolar healing length outside which the dipolar force is greater than the elastic forces. Beyond the soft core, $\mathbf{l}$ is practically uniform; all vorticity in the liquid is thus localized inside the soft core. The continuous vortices have no singular core; however, depending on the external conditions, they may acquire a soft core with localized vorticity. The coherence length in ^{3}He is about two orders of magnitude longer than that in ^{4}He.

A very different type of singular vortex is possible if one takes into account that the total wavefunction for Cooper pairs in

^{3}He-A is the pure product of the orbital and spin parts: $\Psi(\mathbf{r}) = a_{0+}\chi_{10}Y_{1+}$ [see equation (4)], where the dependence on the coordinate **r** of the centre of Cooper-pair mass arises from the **r**-dependencies of the quantization axes: $\mathbf{l}(\mathbf{r})$ and $\mathbf{d}(\mathbf{r})$. The quantization of circulation in 'classical' superfluids is simply a reflection of the fact that the phase of the order parameter, upon circling a closed contour in the superfluid, can only change by $\pm 2\pi m$ because the wavefunction must be single-valued, thus $\Psi(\mathbf{r}) \rightarrow \Psi(\mathbf{r})e^{\pm i2\pi m} = \Psi(\mathbf{r})$. But in ^{3}He-A, we may have for the orbital part of $\Psi(\mathbf{r})$: $Y_{1+}(\mathbf{r}) \rightarrow Y_{1+}(\mathbf{r})e^{\pm i\pi}$, while for the spin part: $\chi_{10}(\mathbf{r}) \rightarrow \chi_{10}(\mathbf{r})e^{\pm i\pi}[\mathbf{d}(\mathbf{r}) \rightarrow -\mathbf{d}(\mathbf{r})]$, resulting in: $\Psi(\mathbf{r}) \rightarrow \Psi(\mathbf{r})e^{\pm 2i\pi} = \Psi(\mathbf{r})$. Hence, half-quantum vortices are possible. However, a half-quantum singularity in the orbital part of $\Psi(\mathbf{r})$ is not separable from a simultaneous singularity in the spin part. Thus in ^{3}He-A we may have an example of topological confinement of singularities that is closely related to similar phenomena in Grand Unified Theories.

Therefore, three topologically different types of vortices[23] can exist in superfluid ^{3}He-A: (1) singular vortices with an odd quantum of circulation, and with topological charge 1; (2) continuous coreless vortices with even integer quanta of circulation, and with topological charge 0; and (3) singular vortices with a half-integer quantum of circulation, and topological charges $\frac{1}{2}$ and $-\frac{1}{2}$. The topological summation laws are: $\frac{1}{2}+\frac{1}{2}=1$, $1+1=0$, and $\frac{1}{2}+1=-\frac{1}{2}$. This means, in particular, that two singular singly quantized vortices can combine to produce a doubly quantized continuous vortex—with consequent annihilation of the singularity.

In ^{3}He-A under rotation, a vortex lattice and a corresponding periodic **l**-vector texture[24] is present: there is a great difference in the quantization conditions for ^{3}He-A and for superfluids with traditional properties. In the latter case, circulation along any closed contour around the vortex is an integer in terms of the circulation quantum. For the vorticity in ^{3}He-A, this is valid only along the boundary of a primitive cell. From energy considerations alone, at angular velocities of the order of 1–3 rad s^{-1}, it follows that, in the so far experimentally relevant magnetic fields of 15–60 mT, the most advantageous structure in the rotating bulk ^{3}He-A is a lattice of singular singly quantized vortices.

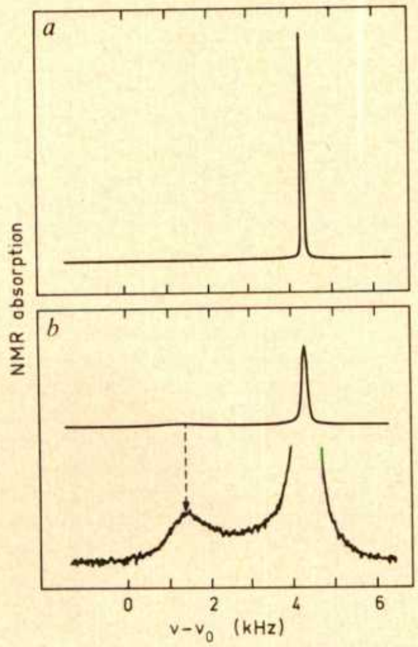


Fig. 4 NMR spectra[11] of stationary and rotating ^{3}He-A. Additional broadening of the main peak, due to rotation, is apparent; the shift of this peak towards higher frequencies by 4.6 kHz from ν_0 is related to the spin-orbit coupling. Arrow points to the satellite absorption peak which appears in the NMR signal during rotation. It is caused by the excitation of spin waves trapped by the continuous vortex texture. Lowest curve is vertically magnified ×50. *a*, $\Omega = 0$; *b*, $\Omega = 1.21$ rad s^{-1}.

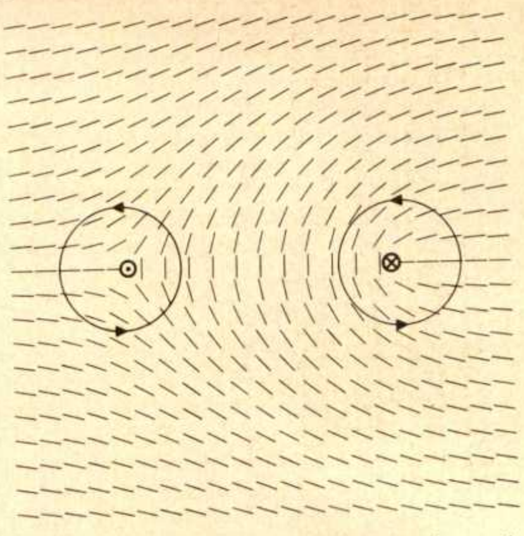


Fig. 5 Texture of the magnetic anisotropy vector **d** around a pair of vortices with half-quantum circulation[28] in ^{3}He-A. Energy of the local disturbance ('soliton') in the uniform **d** texture between the half-quantum pair is proportional to their separation; the pair is thus glued together by the soliton string, as in quark confinement; see Fig. 8*b*. The liquid-crystal-like **d** texture displays defects with opposite Frank indices at the half-quantum vortices: **d** rotates by $+\pi$ and by $-\pi$ on circling the two singular points (⊙ and ⊗).

NMR on vortices

How does one detect and investigate the physical properties of individual vortex lines in ^{3}He? The most versatile tool used so far is the NMR technique, based on the resonance absorption of electromagnetic radiation at the Larmor frequency $\nu_L = \gamma H_0$, corresponding to the precession of nuclear magnetization in the liquid under a constant field H_0; γ is the gyromagnetic ratio of ^{3}He. Note that NMR cannot be used on ^{4}He because the nuclear spin $S = 0$ for this isotope.

Precession of the nuclear spins **S** in an anisotropic superfluid is not free: the directions of the **l** and **d** vectors tend to be locked by the spin-orbit coupling. This produces an additional torque on the spins of the Cooper pairs, and because of this a shift in the NMR absorption line away from γH_0 occurs, see Fig. 4.

In a strong magnetic field, the doubly quantized vortex texture is modified: If **H** is applied parallel to the axis of the rotating vessel, the **d** vector is in the plane perpendicular to $\mathbf{H} \| \mathbf{\Omega}$ because of magnetic anisotropy. Due to the spin-orbit coupling, **l** tends to be parallel to **d**, thus $\mathbf{l} \perp \mathbf{H}$. This results in the broken axisymmetry of the ^{3}He-A vortices in a magnetic field. The distribution of the **l** vector deviates from the homogeneous situation $\mathbf{l} \| \mathbf{d}$ only in the central part of every elementary vortex-lattice cell inside the soft core, where the energy of the inhomogeneity is larger than that due to the spin-orbit coupling. Thus the soft cores, which contain the vorticity, are well separated: at the typical rotational speed of 1 rad s^{-1}, the spacing between the vortices is ~0.3 mm and thus more than an order of magnitude larger than the diameter of the soft core.

The regions of inhomogeneous distribution of **l** act as potential wells for localized standing spin waves. Therefore, a periodic **l**-vector texture leads to the appearance of an additional satellite NMR peak in the absorption spectrum at the frequency of the local vortex mode (Fig. 4). The intensity of this peak is proportional to the number density of vortices, that is to the angular velocity of rotation [see equation (3)]. The calculated intensity of the vortex peak for a lattice of continuous vortices[25–27], each with two circulation quanta, is ~10 times smaller than that for a singular vortex lattice and is in fair agreement with the experi-

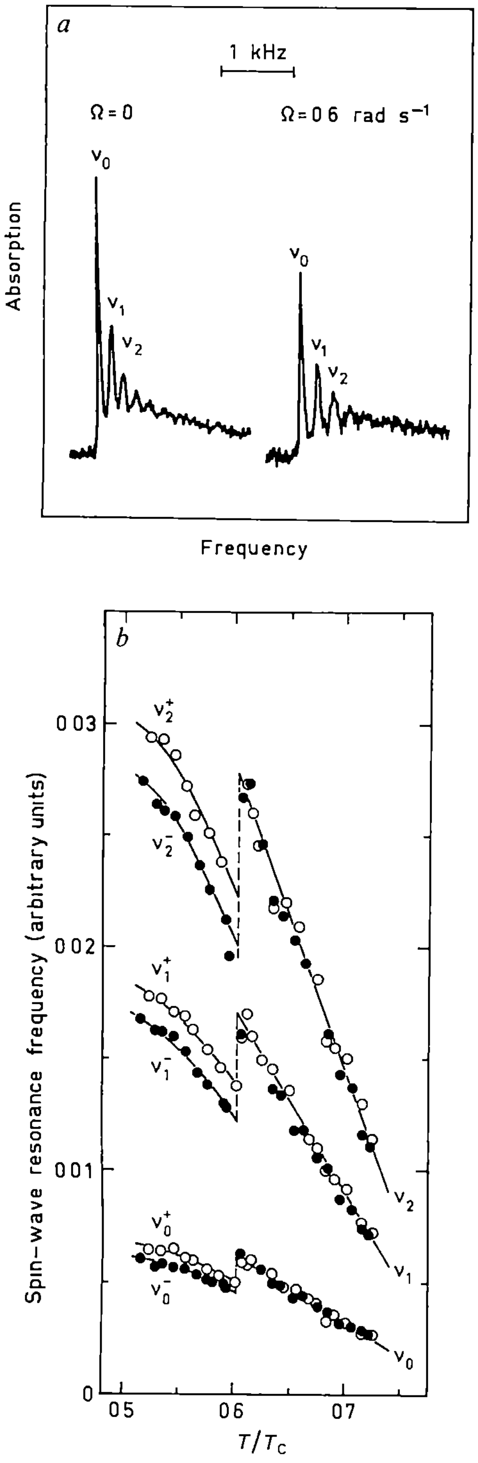


Fig. 6 *a*, Measured NMR absorption spectra[11], at 29.3-bar pressure, for ^{3}He-B as functions of frequency in the stationary case and at $\Omega = 0.6$ rad s^{-1}. The envelope of the observed oscillatory curve corresponds to the 'flare-out' texture[31,32]. The series of almost evenly spaced spin-wave satellites results from the coherent nuclear motion of the Cooper pairs. Separation between the resonant peaks at ν_0, ν_1, ν_2, is larger during rotation than in the stationary liquid. *b*, The spin-wave resonance frequencies ν_0, ν_1, and ν_2, of Fig. 6*a* at $\Omega = 0.6$ rad s^{-1}, as functions of reduced temperature. The data illustrate a gyromagnetic effect[35] due to the magnetic moment of the vortex cores[36,37]: the resonance frequency measured during rotation with **Ω** parallel to the magnetic field **H** (○) differs from that for **Ω** antiparallel to **H** (●). The effect is large at $T < 0.6T_c$ but hardly observed for $T > 0.6T_c$, the vortices below and above this transition temperature at $T_v = 0.6T_c$ have different core structures and, therefore, different intrinsic magnetic moments.

mentally observed NMR satellite intensity[11]. This supports the existence of continuous vortices in ^{3}He-A, instead of the energetically slightly more advantageous singular singly quantized vortices. A possible explanation for this is that in the creation of a singular vortex a high potential barrier has to be overcome. Therefore, we may conjecture that continuous vortices are formed first under rotation and that once they exist in metastable states it becomes difficult to transform one continuous vortex into two singular ones for creating a singular vortex array.

Half-quantum vortices have not been seen experimentally in ^{3}He-A (see, however, below and Figs 2 and 8*b*). These singularities may be observed in the A phase only when the spin and orbital variables can be considered independent. Such a situation exists in narrow gaps between parallel plates, where the energy of the inhomogeneity is larger than the spin-orbit coupling[22,28]. Because of the peculiar topology, the doubly (±) valued **d**-vector field changes sign (∓) on circling a half-quantum vortex line, see Fig. 5. This produces an interesting consequence[29]: the equation for the spin wavefunction bound by the inhomogeneous **d**-vector texture is analogous to the Schrodinger equation for a charged particle in a 'magnetic field', concentrated in the vortex-core tube. This analogue of the Aharonov-Bohm effect also generates spin wave scattering, which should result in significant additional broadening of the main NMR line, possibly providing an experimental signature allowing identification of half-quantum vortices.

Vortices in ^{3}He-B

In 1978, when the decision was made to build a rotating ^{3}He cryostat in Helsinki, it was believed that the vortex lines in superfluid ^{3}He-B would be just like those in superfluid ^{4}He, with no internal structure. In the first NMR experiments[3], however, a spontaneous phase transition in the core was discovered. This immediately implied a complicated vortex core, it is appreciated that vortices in superfluid (^{3}He-A and) ^{3}He-B provide a beautiful application for the concept of broken symmetry[30] in condensed-matter physics. The vortex-core transition has been established to involve a change in the topology of the vortex-core matter between two distinct ways that vorticity can flare out from real to momentum space.

NMR spectroscopy

In ^{3}He-B, the mutual orientation of the quantization axes of L and S is arbitrary in the absence of spin-orbit coupling, the superfluid is thus degenerate with respect to rotations of the spin axis $\hat{\boldsymbol{\zeta}}$ about the orbital axis $\hat{\mathbf{z}}$. Any three-dimensional rotation may be represented by an angle θ around some axis specified by a unit vector $\hat{\mathbf{n}}$. The B-phase states are degenerate with respect to directions of $\hat{\mathbf{n}}$ and, in the absence of spin-orbit coupling, to the value of θ. Moreover, as in superfluid ^{4}He, the ^{3}He-B states are also degenerate with respect to the value of the overall phase Φ of the order parameter. This means that the energy of the homogeneous liquid does not depend on $\hat{\mathbf{n}}$, θ or Φ. In inhomogeneous ^{3}He-B, for example, when vortices are present, the energy depends on the gradients of $\hat{\mathbf{n}}$, θ and Φ.

As in ^{4}He, the superflow velocity in ^{3}He-B is determined by the gradient of the phase $\mathbf{v}_s = (\hbar/2m_3)\nabla\Phi$ and the circulation of superflow is quantized. Therefore, there are no coreless vortices in ^{3}He-B and the superfluidity of this phase is in several respects analogous with the properties of superfluid ^{4}He. In particular, a two-dimensional lattice of singly quantized vortices is formed under rotation.

In ^{3}He-B, the order parameter and the texture are fully determined by the $\hat{\mathbf{n}}$ vector, which is analogous to the directrix vector describing the orientation of molecules in liquid crystals. Because the orientational force exerted by an external magnetic field on the nearly isotropic ^{3}He-B is so small, the magnetic healing length ξ_H is long, effects due to walls are thus very important in ^{3}He-B and they give rise to $\hat{\mathbf{n}}$ textures[31,32]. Inside the vortex cores, however, the magnetic anisotropy energy is of

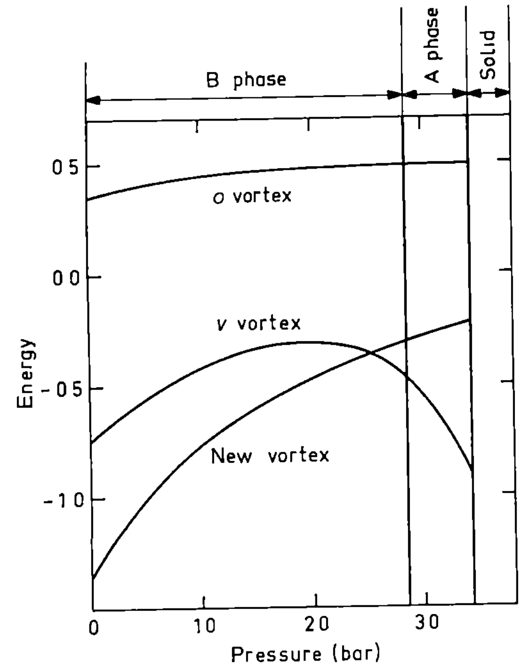


Fig. 7 Energies of the axisymmetric *o* and *v* vortices, and the nonaxisymmetric new vortex, as functions of pressure according to numerical calculations[41]. Pressures are given on a theoretical scale: for example, the tricritical point (see Fig. 2), actually at 21 bar, is at 28.5 bar.

the same order of magnitude as in ^{3}He-A, therefore, vortices have a strong influence on the $\hat{\mathbf{n}}$ texture. The $\mathbf{v}_s$ field around the vortex also has an orientating effect on $\mathbf{n}$ (ref. 33).

The most powerful experimental tool for observing vortices in ^{3}He-B is the NMR technique, as in ^{3}He-A. The precessing motion of the nuclear spins is coupled with the order-parameter distribution, determined by the spatial variations of $\hat{\mathbf{n}}$, θ and Φ. The energy minimum for spin-orbit coupling fixes θ to the 'magic' angle $\theta_0 = 104°$. Another consequence of the coupling is the appearance of a small anisotropy, caused by the magnetic energy term which orients $\hat{\mathbf{n}}$ along $\mathbf{H}$. In the presence of an axial magnetic field, as a result of boundary conditions on $\hat{\mathbf{n}}$ and the magnetic anisotropy, the equilibrium orientation only occurs in the bulk liquid at distances $> \xi_H[\mathrm{cm}] \simeq 4(1 - T/T_c)/H[\mathrm{mT}]$ from the walls.

In a cylindrical container, with an axially oriented $\mathbf{H}$, a spiral-like 'flare-out' texture results[31,32]. As in ^{3}He-A, the $\hat{\mathbf{n}}$-texture gives rise to observable localized standing spin waves. These stationary coherent nuclear spin modes are confined by the flare-out texture (Fig. 6*a*). The frequencies of these modes depend on the spatial variation of $\hat{\mathbf{n}}$. With the use of NMR it is thus possible to register changes in $\hat{\mathbf{n}}$. Consequently, if the texture of the liquid is altered by rotation, this can be observed. The influence of rotation on $\hat{\mathbf{n}}$ is due to those vortex-core regions where the balance between the equal-weight projections of $S_\zeta = \pm 1, 0$ and $L_z = \pm 1, 0$ is broken. Because of this, the cores of vortices in ^{3}He-B possess an anisotropic magnetic susceptibility[33], which depends on the orientation of $\hat{\mathbf{n}}$ and on the direction of $\mathbf{\Omega}$.

The appearance of an additional energy of magnetic anisotropy due to the vortex cores suggests the possibility of studying the vortices by the measurement of the Ω dependence of the frequencies of the spin-wave modes in the flare-out texture, see Fig. 6. The signature of vortices is also in the shift of the NMR signal[34] in a field $\mathbf{H}$ tilted with respect to the direction of the axis of rotation: the shift is zero without the vortices. In addition, a gyromagnetic effect[35] has been observed, which was manifested by the fact that the cores of the ^{3}He-B vortices possess a magnetic moment[36-39], for further discussion, see Fig. 6*b* and below.

Vortex structures

In superfluid ^{3}He, the hard vortex-core size is in the Ginzburg-Landau (GL) regime near T_c of the order of several coherence lengths, $\xi_{GL} \simeq 10\ (1 - T/T_c)^{-1/2}$ nm, which is the temperature-dependent Cooper-pair extent. Because of the large size of vortices and the complicated nature of ordering it is not necessary to have normal liquid inside the vortex cores in ^{3}He-B, instead, other combinations of different p-wave pairing states may nucleate. The structure of the core can be determined by numerical minimization of the vortex energy.

In the cylindrical coordinate frame[36], the order-parameter amplitudes $a_{\mu\nu}(\mathbf{r})$ in equation (4) may be written for a singly quantized vortex as

$$a_{\mu\nu}(\mathbf{r}) = \begin{matrix} C_{++}\,e^{-i\phi} & C_{+0} & C_{+-}\,e^{i\phi} \\ C_{0+} & C_{00}\,e^{i\phi} & C_{0-}\,e^{2i\phi} \\ C_{-+}\,e^{i\phi} & C_{-0}\,e^{2i\phi} & C_{--}\,e^{3i\phi} \end{matrix} \qquad (5)$$

For an axisymmetric vortex, the above amplitudes $C_{\mu\nu}$ are only functions of the radial distance r from the vortex axis and independent of the azimuthal angle ϕ. The core structures have been classified using symmetry arguments[37,38]: it was found that an axisymmetric vortex in ^{3}He-B, with unit quantum of circulation, may be in five different states, distinguished by the internal discrete symmetry of the vortex core. These states are labelled *o*, *u*, *v*, *w* and *uvw*, in all of them the vortex core possesses a magnetic moment. The most symmetric *o* vortex, first discussed by Ohmi *et al.*[36], has a normal core with superfluidity destroyed on the axis, as in ^{4}He. In the *u*, *v* and *w* vortices the internal symmetry is partially destroyed, whereas in the *uvw* vortex it is completely broken.

Energetically the most favourable of the different possible axisymmetric vortices is the *v* vortex with broken parity[37,38], in this case, all the nine functions $C_{\mu\nu}$ in equation (5) are real. Although the B-phase pairing amplitudes C_{+-}, C_{00}, and C_{-+} also in this vortex must tend to zero on the vortex axis, because of broken parity, this vortex, nevertheless, has a superfluid core: the two components without phase factors in equation (5), the *A* phase C_{0+} and a ferromagnetic so-called β phase C_{+0}, are not prohibited by any centrifugal barrier at the vortex axis. This results in the core which practically consists of the A phase, as it is energetically more favourable than the β phase. These results were later independently confirmed by other workers[39-41].

Because of the ferromagnetic superfluid in the core, a large spontaneous vortex magnetization $\mathbf{M}$ results, its direction is determined by the directions of $\mathbf{\Omega}$ and $\hat{\mathbf{n}}$. This vortex magnetization, created by rotation, gives rise to a new free-energy term, $-\mathbf{H}\cdot\mathbf{M}$, which is odd in both $\mathbf{\Omega}$ and $\mathbf{H}$ and which explains the gyromagnetic effect shown in Fig. 6*b*. A change in the sense of rotation or in the direction of the magnetic field changes the texture of the liquid. Therefore, $\mathbf{M}$ could be measured[35] directly by NMR spectroscopy even though its value in ^{3}He-B rotating at $\Omega = 1$ rad s^{-1} is tiny: only of the order of 10^{-11} nuclear Bohr magnetons per ^{3}He atom.

The *v* vortex possesses a spontaneous electric polarization, also caused by the violation of parity in the core[38]. The *w* vortex would also show striking properties: it has a complex superfluid core and it displays spontaneous superflow along the vortex axis. The *uvw* vortex shares the properties of both the *v* and *w* vortices.

Vortex-core transition in ^{3}He-B

Perhaps the most interesting feature signalling a nontrivial vortex-core structure in ^{3}He-B is the phase change, which takes place when crossing a transition line while the liquid is rotating, see Fig. 2. This was observed experimentally as abrupt discontinuities in the spacing of standing spin waves[3] and as a jump in the vortex-core magnetization[35] (Fig. 6*b*) and as a jump in the critical flow velocity[14,15]. These changes are independent of the angular velocity of rotation, that is, the density of vortices,

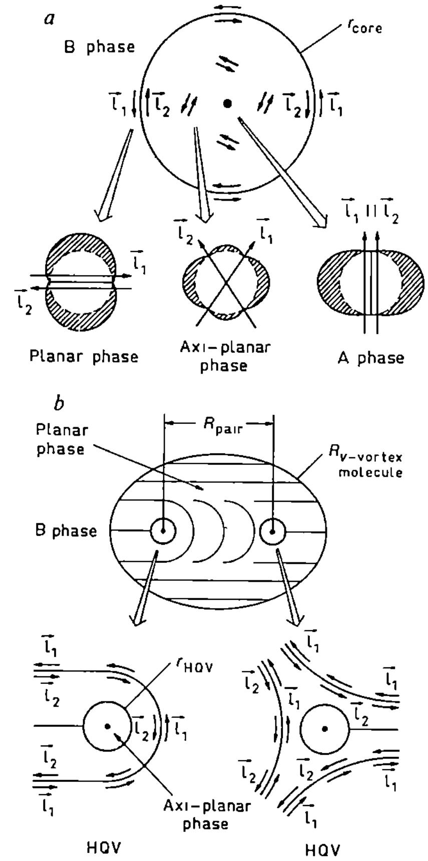


Fig 8 Topology of the energy gap (hatched areas) on the Fermi sphere in the cross-sectional plane of *v* vortices in ³He-B at high and low pressures (see Fig 2) *a*, In the axisymmetric *v* vortex[38], point vortices first appear in the 'planar-like' phase at $r = r_{core}$ For $r < r_{core}$ vorticity is supported by continuous winding, in the 'axi-planar-like' phase, of the nodes for quasiparticles with spin up ($\mathbf{l}_1$) and down ($\mathbf{l}_2$), changing from antiparallel at $r = r_{core}$ to parallel at $r = 0$ with 'A-phase-like' core *b*, The nonaxisymmetric *v*-vortex 'molecule' The two centres are both a half-quantum vortex[49] (HQV) confined in the core (see Fig 5) Each HQV is accompanied by a disclination in the l-vector field with half-integer winding number, reversing $\mathbf{l}_1$ into $\mathbf{l}_2$ on circling either HQV The nodes lie in the cross-sectional plane ('planar-like' phase), except inside inner cores of radius r_{HQV} where they flare out of the plane and deviate from each other (in the 'axi-planar-like' phase) This 'bag' model of the core 'molecule' resembles hadrons with the HQVs corresponding to quarks

implying that the transition is not affected by interactions between vortices Rather, the transition is attributed to a change in the core of isolated vortex lines

Much work has been devoted to identifying the vortices associated with this unexpected transition First, it was contemplated[37,42] that one might understand the transition to be one between two of the five axisymmetric vortices It was found, however, that the *v* vortex provides the absolute energy minimum among the axisymmetric vortices[38] Also, it is stable against small nonaxial perturbations[43] Salomaa and Volovik[44] suggested that the mechanism of the vortex-core transition in ³He-B is breaking of axisymmetry Nonaxisymmetric vortex structures were also considered by Theodorakis and Fetter[45]

Recently, it has been reported[41,46,47] that a nonaxisymmetric vortex has a lower energy than the axisymmetric *v* vortex at low pressures Thuneberg[41] was the first to perform the full two-dimensional energy minimization by solving numerically the Ginzburg-Landau differential equations, his results are illustrated in Fig 7, whereas Volovik and Salomaa[46,47] only carried out investigations within sectors of definite symmetry At all pressures the energy of the axisymmetric *o* vortex is clearly higher than the energies of the axisymmetric *v* vortex[37] or the nonaxisymmetric 'double core' new vortex with two points of local maxima in the energy density At low pressures, the new vortex has the lowest energy With an increase of pressure, the energy difference between the new vortex and the axisymmetric *v* vortex decreases and their energies are equal at ~3 bar below the tricritical pressure, this is in excellent agreement with the experimental data of Fig 2 Thuneberg's calculations show that the axisymmetric *v* vortex becomes stable at higher pressures

Volovik and Salomaa[46,47], on the basis of their systematic study[38] of all the discrete broken symmetries, confirmed Thuneberg's[41] calculation and identified the new vortex as the nonaxisymmetric *v* vortex of their classification scheme At present it seems that in ³He-B the nonaxisymmetric *v* vortex is stable at low pressures and the axisymmetric *v* vortex at high pressures (Figs 2 and 7) Volovik and Salomaa find that the mechanism of restructuring the superfluid core in the vortex transition is independent of any model-dependent details the first-order vortex-core transition is triggered by the proximity of the *A* phase alone, and driven by the broken parity of the *v* vortex, combined with superfluid condensation-energy considerations

Point vortices on the Fermi surface

What is the actual physical meaning of the vortex-core transition what happens to superfluid vorticity? To answer these questions, it is crucial to appreciate that vortices in ³He can couple the real **r** and the momentum **p** spaces superfluid-core vortices can escape the real-space singularity on the vortex axis by flaring out into the momentum space, this process is associated with nodes in the energy gap on the Fermi surface within the vortex core[48] We denote by $\mathbf{l}_1$ and $\mathbf{l}_2$ the directions of the nodes for the two energy gaps, $\Delta_1(\hat{\mathbf{p}})$ and $\Delta_2(\hat{\mathbf{p}})$, for the two possible ³He spin projections In both *v* vortices, there appear two pairs of such nodes These objects are topologically interesting point vortices ('boojums') in the **p**-space, around which the phase Φ_i of $\Delta_i = |\Delta_i(\hat{\mathbf{p}})|\, e^{i\Phi_i(\hat{\mathbf{p}})}$ (for $i = 1, 2$) changes by $\pm 2\pi$, therefore, the singularity ($v_s \propto 1/r$) can be resolved without breaking superfluidity on the vortex axis

Volovik and Mineev[48] were the first to propose that the core transition in ³He-B involves a change in the topology of the vortex-core matter, however, the transition proved more interesting[49] than that originally suggested To escape the singularity in the axisymmetric *v* vortex[38] (Fig 8*a*), $\mathbf{l}_1$ and $\mathbf{l}_2$ each cover half the Fermi sphere unwinding unit quantum of circulation thus results according to $m = \frac{1}{2} + \frac{1}{2} = 1$, this is associated with the simultaneous spatial variation of each orbital anisotropy axis, $\mathbf{l}_1$ and $\mathbf{l}_2$, just like vorticity can be supported by the continuous winding of the order parameter in the **l** textures of ³He-A

For the nonaxisymmetric *v* vortex, the topological scenario for flaring-out of the singularity is essentially different[49] The composite core 'molecule' is illustrated in Fig 8*b* It contains two half-quantum vortices, 'quarks'[47], coupled by 'strings' disclinations in both $\mathbf{l}_1$ and $\mathbf{l}_2$, closely resembling the **d**-field disclination in Fig 5 One may here verify that each point vortex covers one quarter of the Fermi surface in both of the half-quantum vortices $m = (\frac{1}{4} + \frac{1}{4}) + (\frac{1}{4} + \frac{1}{4})$, hence the two 'quarks' do represent objects with fractional topological charge Consequently, Fig 2 identifies the vortices involved in the core transition, whose first-order character signals the topological bifurcation in the distribution of vorticity into the momentum space

Conclusions

The varied vortex-core structures in the superfluid phases of ³He lead to the possibility of several phase transitions inside the cores These can occur in the hard vortex core, but also in the much more extended soft-core region In ³He-B, so far only one of many possible transitions has been found in the NMR spectra[3,15,35], the nature of this vortex-core transition is understood theoretically[49] The search and study of these new physical phenomena will be the subject of future theoretical and experimental investigations

To date, these studies, principally with use of the NMR technique, have concentrated on the static properties of vortices In the future, problems related to the dynamics of vortices need to be tackled nucleation, the mechanisms for transitions—possibly involving monopoles—to lower-energy states, and the formation and properties of vortex rings In this research, ions[50] have already been used and experiments involving ultrasound are being planned

The physics of vortex phenomena in ^{3}He is interesting from a much broader point of view, owing to the different mechanisms for topological confinement, exemplified by the generation of the **d**-vector soliton (string) that glues together the pairs of half-quantum vortex disclinations in ^{3}He-A, as in quarks There are many conceptual connections to quantum field theories, and close analogies with defects in liquid crystals[51] The structure of the vortex core in the B phase is intimately related to the topology of the point vortices on the Fermi surface[48] The novel dimerization mechanism in flaring-out of vorticity for the non-axisymmetric v vortex molecule suggests the identity of the vortex-core transition in ^{3}He-B the nonaxisymmetric v vortices are topologically confined pairs of half-quantum vortices, analogous to quarks inside nucleons Provided that experiments will confirm this identification[49], vortices with half-integer circulation have first been found in rotating superfluid ^{3}He-B

It is commonly believed that the neutron mass in pulsars undergoes a BCS-type transition to form a superfluid Fermi liquid in a $^3P_2$ state[52] Normal-core vortices with spontaneous magnetization have been discussed for neutron stars[53,54], and the analogue experiments on the superfluidity of pulsars[55] have been conducted on rotating superfluid ^{4}He Being a Fermi liquid, also paired in a p-wave state, superfluid ^{3}He should provide a better substance for such model experiments It is likely that, just as in superfluid ^{3}He, vortices in pulsars also display a superfluid core This possibility could introduce a substantial revision in present understanding of heat transport in neutron stars and thereby in our picture of the latter stages of supernovae

An interesting experiment was recently suggested by Zurek[56], the proposal is to carry out a rapid isothermal expansion of liquid ^{4}He through the superfluid transition, whereby vorticity should be generated at the interfaces of the superfluid condensation fronts Such vortex lines are proposed to be analogous to the cosmic strings[57,58] that are believed to have nucleated in the phase transitions of the early Universe An analogue model experiment on superfluid ^{3}He (here presumably by a rapid isothermal compression through the superfluid transition) would be more difficult to carry out However, this is potentially more interesting because the 'vacuum' states of superfluid ^{3}He, with the most complicated known broken symmetry in condensed matter, provide a more relevant analogue system for Grand Unified Theories Here the vortex density can also be directly measured by the NMR technique This idea has also been proposed for liquid crystals (M Yu Khlopov and S Obukhov, unpublished data), where the creation of disclination lines at the phase transition may be seen directly with a microscope

Vortices in heavy-fermion superconductors remain largely unexplored Provided that these materials[59] do exemplify p-wave or other unconventional pairing, it is possible that some of the conceptual generalizations which have been necessary to make progress in understanding experimental results on vortices in superfluid ^{3}He will also prove useful in the context of heavy-fermion superconductors

In conclusion, vortices in superfluid ^{3}He, although highly interesting for their own sake also seem to be important in a broader context beyond low-temperature physics, for more detailed expositions of the topic, see refs 34, 43 and 60 (review)

We thank P Berglund, Yu M Bun'kov, A L Fetter, I A Fomin, A D Gongadze, P J Hakonen, O T Ikkala, S T Islander, G A Kharadze, M Krusius, K Maki, N D Mermin, T Ohmi, R E Packard, J P Pekola, J D Reppy, J A Sauls, N Schopohl, J T Simola, the late J S Tsakadze and especially G E Volovik for collaboration at various stages of the work reported here, and for helpful comments This research has resulted from the ROTA project, supported by the USSR Academy of Sciences, by the Commission for Scientific and Technical Collaboration between Finland and the USSR, and by the Academy of Finland

1 Tilley, D R & Tilley, J *Superfluidity and Superconductivity* (Adam Hilger, Bristol, 1985)
2 Andronikashvili E L & Mamaladze, Yu G *Rev Mod Phys* **38**, 567–625 (1966)
3 Ikkala, O T *et al Pis ma Zh eksp teor Fiz* **35**, 338–341 (1982), *JEPT Lett* **35**, 416–420 (1982)
4 Feynman, R P in *Progress in Low Temperature Physics* Vol I, Chap 2 (ed Gorter, C J) (North Holland, Amsterdam, 1964)
5 Putterman, S J *Superfluid Hydrodynamics* (North Holland, Amsterdam, 1974)
6 Tinkham, M *Introduction to Superconductivity* (McGraw-Hill, New York, 1975)
7 Leggett, A J *Rev Mod Phys* **47**, 331–414 (1975)
8 Wheatley, J C *Rev Mod Phys* **47**, 415–470 (1975)
9 Lee, D M & Richardson, R C in *The Physics of Liquid and Solid Helium* Vol II, Chap 4 (eds Bennemann, K H & Ketterson, J B) (Wiley, New York, 1978)
10 Hakonen, P J *et al Cryogenics* **23**, 243–250 (1983)
11 Hakonen, P J , Ikkala, O T , Islander, S T , Lounasmaa, O V & Volovik, G E *J Low Temp Phys* **53**, 425–476 (1983)
12 Gammel, P L , Hall, H E & Reppy, J D *Phys Rev Lett* **52**, 121–124 (1984)
13 Gammel, P L , Ho, T L & Reppy, J D *Phys Rev Lett* **55**, 2708–2711 (1985)
14 Pekola, J P , Simola, J T , Nummila, K K , Lounasmaa, O V & Packard, R E *Phys Rev Lett* **53**, 70–73 (1984)
15 Pekola, J P *et cl Phys Rev Lett* **53**, 584–587 (1984)
16 Ambegaokar, V , deGennes, P G & Rainer, D *Phys Rev* A **9**, 2676–2685 (1974), erratum *Phys Rev* A **12**, 345 (1975)
17 Mermin, N D & Ho, T L *Phys Rev Lett* **36**, 594–597 (1976), erratum *Phys Rev Lett* **36**, 832 (1976)
18 Volovik, G E *Usp Fiz Nauk* **143**, 73–109 (1984), *Soviet Phys Usp* **27**, 363–384 (1984)
19 Chechetkin, V R *Zh eksp teor Fiz* **71**, 1463–1474 (1976), *Soviet Phys—JETP* **44**, 766–772 (1976)
20 Anderson, P W & Toulouse, G *Phys Rev Lett* **38**, 508–511 (1977)
21 Volovik, G E & Mineev, V P *Pis ma Zh eksp teor Fiz* **24**, 605–608 (1976), *JETP Lett* **24**, 561–563 (1976)
22 Cross, M C & Brinkman, W F *J Low Temp Phys* **27**, 683–686 (1977)
23 Volovik, G E & Mineev, V P *Zh eksp teor Fiz* **72**, 2256–2274 (1977), *Soviet Phys—JETP* **45**, 1186–1196 (1977)
24 Volovik, G E & Kopnin, N B *Pis ma Zh eksp teor Fiz* **25**, 26–28 (1977), *JETP Lett* **25**, 22–24 (1977)
25 Seppälä, H K & Volovik, G E *J Low Temp Phys* **51**, 279–290 (1983)
26 Seppalä, H K *et al Phys Rev Lett* **52**, 1802–1805 (1984)
27 Zotos, X & Maki K *Phys Rev* B **30**, 145–151 (1984)
28 Volovik, G E & Salomaa, M M *Zh eksp teor Fiz* **88**, 1656–1663 (1985), *Soviet Phys—JETP* **61**, 986–990 (1985)
29 Salommaa, M M & Volovik, G E *Phys Rev Lett* **56**, 1313 (1986)
30 Anderson, P W *Basic Notions of Condensed Matter Physics* (Benjamin, Menlo Park, California 1984)
31 Smith, H , Brinkman, W F & Engelsberg, S *Phys Rev* B **15**, 199–213 (1977)
32 Maki, K & Nakahara, M *Phys Rev* B **27**, 4181–4185 (1983)
33 Gongadze, A D , Gurgenishvili, G E & Kharadze, G A *Fiz Nizk Temp* **7**, 821–825 (1981), *Soviet J Low Temp Phys* **7**, 397–398 (1981)
34 Bun'kov, Yu M , Gurgenishvili, G E , Krusius, M & Kharadze, G E *Usp Fiz Nauk* **144**, 141–168 (1984), *Soviet Phys Usp* **27**, 731–747 (1984)
35 Hakonen, P J *et al Phys Rev Lett* **51**, 1362–1365 (1983)
36 Ohmi, T , Tsuneto, T & Fujita, T *Prog theor Phys* **70**, 647–653 (1983)
37 Salomaa, M M & Volovik, G E *Phys Rev Lett* **51**, 2040–2043 (1983)
38 Salomaa, M M & Volovik, G E *Phys Rev* B **31**, 203–227 (1985)
39 Passvogel, T , Tewordt, L & Schopohl, N *J Low Temp Phys* **56**, 383–414 (1984)
40 Hasegawa, Y *Prog theor Phys* **73**, 1258–1260 (1985)
41 Thuneberg, E V *Phys Rev Lett* **56**, 359–362 (1986)
42 Fetter, A L *Phys Rev* B **31**, 7012–7017 (1985)
43 Fetter, A L in *Progress in Low Temperature Physics* (ed Brewer, D F) Vol X (North-Holland, Amsterdam, in the press)
44 Salomaa, M M & Volovik, G E *Am Inst Phys Proc* No 103, 210–228 (1983)
45 Theodorakis, S & Fetter, A L *J Low Temp Phys* **52**, 559–591 (1983)
46 Volovik, G E & Salomaa, M M *Pis ma Zh eksp teor Fiz* **42**, 421–424 (1985), *JETP Lett* **42**, 521–524 (1985)
47 Salomaa, M M & Volovik, G E *Phys Rev Lett* **56**, 363–366 (1986)
48 Volovik, G E & Mineev, V P *Zh eksp teor Fiz* **83**, 1025–1037 (1982), *Soviet Phys—JETP* **56**, 579–586 (1982)
49 Salomaa, M M & Volovik, G E *Europhys Lett* **2**, 781–788 (1986)
50 Simola, J T *et al Phys Rev Lett* **57**, 1923–1926 (1986)
51 Kleman, M *Points, Lines and Walls in Liquid Crystals, Magnetic Systems and Various Ordered Media* (Wiley, New York, 1983)
52 Pines, D & Alpar, M A *Nature* **316**, 27–32 (1985)
53 Muzikar, P , Sauls, J A & Serene, J W *Phys Rev* D **21**, 1494–1502 (1980)
54 Sauls, J A , Stein, D L & Serene, J W *Phys Rev* D **25**, 967–975 (1982)
55 Tsakadze, J S & Tsakadze, S J *J Low Temp Phys* **39**, 649–688 (1980)
56 Zurek, W H *Nature* **317**, 505–508 (1985)
57 Zel dovich, Ya B , Kobzarev, I Yu & Okun, L B *Zh eksp teor Fiz* **67**, 3–11 (1974), *Soviet Phys—JETP* **40**, 1–5 (1975)
58 Vilenkin, A *Phys Rep* **121**, 263–315 (1985)
59 Fisk, Z , Ott, H R , Rice, T M & Smith, J L *Nature* **320**, 124–129 (1986)
60 Salomaa, M M & Volovik, G E *Rev Mod Phys* (submitted)

Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes

Claudio W. Pikielny*, Brian C. Rymond† & Michael Rosbash†

* Department of Biochemistry and † Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA

Three splicing complexes formed with a yeast pre-messenger RNA during in vitro *splicing can be resolved by non-denaturing gel electrophoresis after incubation in the presence of non-specific competitor RNA The time course of the appearance of these complexes and their composition suggest that they represent an ordered pathway of splicing complex assembly*

THE detailed analysis of RNAs generated during pre-mRNA splicing *in vitro*[1-5] and *in vivo*[6-8] has led to the definition of a two-step splicing pathway in mammalian systems as well as in the yeast *Saccharomyces cerevisiae* (for recent reviews see refs 9, 10) In the first step, the 5′ splice junction is cleaved, generating two molecules, the exon 1, and the intron-exon 2, in a lariat configuration, due to a 2′-5′ linkage between the 5′ end of the intron and the 2′-OH of an adenosine in the intron (which in *Saccharomyces cerevisiae* is the last A of the TACTAAC box) In the second step the two exons are ligated and the intron is excised as a lariat The two-step pathway appears similar for *S cerevisiae* and mammalian systems but differences in the *cis*-acting signals and the effects of mutations in these *cis*-acting sequences, indicate that some aspects of pre-mRNA splicing are different[11-13]

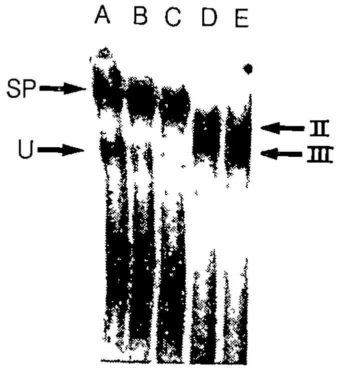


Fig. 1 Effect of non-specific competitor RNA on the electrophoretic resolution of splicing complexes A splicing reaction was carried out as described[4,5,11] with recombinant plasmids transcribed *in vitro* with SP6 polymerase[29] and containing rp51A sequences[13] The reaction was incubated for 20 min at 25 °C, stopped by the addition of Q buffer (400 mM KCl, 2 mM $Mg(OAc)_2$, 20 mM EDTA and 100 mM hydroxyethyl piperazine ethane sulphonic acid HEPES-NH_4OH (pH 7 5)), and separated into 10 μl aliquots which were analysed by electrophoresis on a non-denaturing agarose-acrylamide composite gel[13] Just before loading, each 10 μl sample was incubated on ice for 10 min in the presence of a preparation of total yeast RNA[30] at a concentration of lane A, 0 μg μl^{-1}, B, 0 1 μg μl^{-1}, C, 0 2 μg μl^{-1}, D, 0 4 μg μl^{-1}, E, 1 μg μl^{-1} The optimal concentration of non-specific RNA is ~1 μg μl^{-1}, higher concentrations did not improve the resolution but caused streaking and aggregation (not shown) Other nucleic acids have been tested as competitors, but neither transfer RNA nor DNA improve the resolution of splicing complexes *Xenopus laevis* total RNA was just as effective as was *Saccharomyces cerevisiae* total RNA The complexes observed, in the different lanes, labelled SP, U, II and III

Splicing occurs in very large RNP (ribonucleoprotein) complexes[14-16] In mammalian systems these splicing complexes, also called spliceosomes, contain snRNPs (small nuclear ribonucleoproteins[15,17]) which are required for the splicing reaction[9,10] Point mutations in either of the two sequences essential for the splicing of yeast pre-mRNAs (the 5′ splice junction and the TACTAAC box) inhibit spliceosome formation[14,18] It has only recently been possible to show that specific trimethyl-capped RNAs (for a review of small nuclear RNAs, see ref 19) are present in yeast spliceosomes[13] and that at least some of these are essential for the splicing reaction[19]

We recently used non-denaturing polyacrylamide gels to analyse yeast RNPs formed during *in vitro* splicing[13] Of the two RNPs visible, the first (called SP for spliceosome) shows all the characteristics of a splicing complex or spliceosome its formation is dependent on (1) time, (2) ATP and (3) an intact TACTAAC box and 5′ splice junction, it also contains the products of the first step of the splicing reaction (exon 1 and intron-exon 2 lariat intermediate) and a product of the second step, the excised intron in a lariat configuration SP band formation precedes any detectable cleavage or ligation events, reaching a maximal level shortly after the addition of ATP The second complex (named U for unspecific) behaves as a nonspecific RNP by all the above criteria These data imply that only the SP complex is a spliceosome and that its formation is an early and probably essential step in the splicing of pre-mRNA

Competitor RNA improves resolution

The addition of non-specific competitor RNA to a splicing reaction before electrophoresis increases the mobility of both the SP and the U complexes, presumably because loosely bound factors are lost (Fig 1) The U band is readily dissociated with increasing concentrations of competitor RNA, as expected of a nonspecific complex The SP band is less affected, implying that some interactions in this complex are resistant to the addition of competitor RNA The SP band was resolved into two major bands of higher mobility, suggesting that in the absence of competitor RNA the SP band is heterogeneous and contains at least two different splicing complexes The two bands visible in this experiment are bands II and III, the only ones normally detectable after a 20 min incubation (Fig 1)

Three splicing complexes

A time course shows that three distinct complexes, called I, II and III in order of increasing electrophoretic mobility, are formed (Fig 2, w t) In this typical experiment, substantial amounts of bands I and III were visible after 2 min of incubation, the concentrations were maximal at 5 min and decreased thereafter Substantial amounts of bands I and III are formed after 30 s of incubation (data not shown), well before the formation of splicing intermediates and products can be detected (see ref 13, for example) Band II is not visible until 5 min of incubation after which it continues to increase in intensity The first appearance of band II coincides with that of the intermediates of the reaction The order of appearance of the three complexes suggests that two early complexes (bands I and III) may be intermediates in a pathway leading to the formation of the late complex (band II)

Complex formation with point mutant

Analysis of a strong point mutation at the TACTAAC box (the 3'-III mutation, TACTACC[21]) supports and extends this idea As expected from *in vivo* studies[21], incubation of an RNA substrate containing this mutation in an *in vitro* splicing reaction, results in no detectable intermediates or products (data not shown) Surprisingly, however, there was a substantial accumulation of band III (Fig 2, 3'-III, note that this portion of the gel reads right to left), which formed more slowly than with a wild-type substrate RNA No band I or band II was detected The data suggest that the 3'-III mutant is blocked after formation of the early complex (band III) and hence that band III is the initial complex and leads to the subsequent formation of bands I and II with a wild-type substrate

Distribution of labelled RNAs

To characterize these splicing complexes further, we analysed the RNA distribution across RNP gels (Fig 3) Precursor RNA has a broad distribution with a clear maximum which coincides with band III and extends into band II and the region corresponding to band I (band I is indistinct at incubation times of 20 min) The products of the first step, the intron-exon 2 lariat and exon 1, are found almost exclusively in the band II fractions The intron produced in the second step is distributed over bands II and III, in Fig 3 it is more prominent in the band III region A very low level of the exon 1-exon 2 RNA is detectable in band II

We used a smaller RNA to determine the influence of substrate size on the migration of the different splicing complexes and to improve resolution The rp51A δ2 RNA is only 247 nucleotides (nt) long (as opposed to 800 nt for rp51A) but is spliced accurately and efficiently *in vivo*[22] and *in vitro*[11] The smaller substrate gives rise to a similar pattern of bands, the early forming complexes are designated bands III and I and the later complex is named band II (Fig 4*a*) All three of these complexes show much higher mobility than the corresponding complexes formed with the larger substrate There is also a fourth complex of higher mobility, not visible at early times (band IV) The RNAs present in each band from the 20 min time point were analysed on a denaturing polyacrylamide-urea gel (Fig 4*b*), the distribution of radioactive RNA species between the bands is similar to that found for the larger transcripts (Fig 4*b* and data not shown), confirming the correspondence of the complexes formed with the two transcripts Band I contains mostly pre-mRNA and band III contains only pre-mRNA, whereas band II contains substantial amounts of both products of the first step, the intron-exon 2 lariat and exon 1 Bands I and II are not completely separated, but it can be seen that band I contains mostly pre-mRNA and that the products of the first step contribute >50% of the radioactivity in band II Band IV contains mostly intron-exon 2 lariat, some excised intron and pre-mRNA, but no detectable exon 1 (compare lanes II and IV) It probably results from the partial dissociation of band II and cannot be an active complex because it lacks exon 1

snRNAs associated with each complex

We have recently shown through partial purification of *in vitro* generated splicing complexes (or spliceosomes) by two different methods (non-denaturing electrophoresis and oligo(dT)-cellulose chromatography of spliceosomes formed on a polyadenylated substrate) that at least three different trimethyl cap-bearing snRNAs, 160, 185 and 215 nt long are associated with

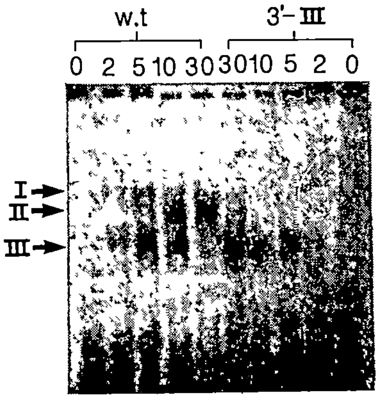


Fig 2 Time course of formation of splicing complexes Splicing reactions were carried out for the times indicated above each lane (in minutes), quenched and incubated in the presence of 1 μg μl^{-1} of total yeast RNA for 10 min on ice just before loading as in Fig 1 (Note that, for the mutant transcript, the time points read from right to left) w t, splicing complexes formed with the wild-type rp51A transcript 3'-III, splicing complexes formed with a transcript indentical to the one in *a* except for an A to C transversion at the last A of the TACTAAC box[21] The different bands observed are labelled I, II and III

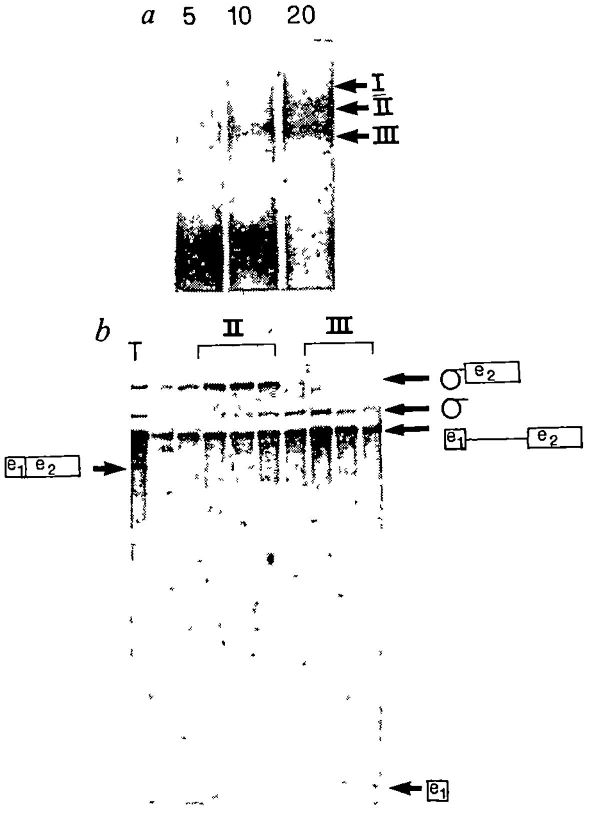


Fig 3 Distribution of pre-mRNA, intermediates and products among splicing complexes formed with a long substrate *a*, Splicing reactions with saturating amounts of substrate (1 ng μl^{-1}) were incubated for 5, 10 and 20 min as indicated and resolved on a non-denaturing gel after incubation with total yeast RNA The gel was then electroblotted to a DEAE membrane, which was then autoradiographed[13] *b*, The lane corresponding to the 20-min time point was cut into 2-mm slices, the RNA eluted from each slice and analysed on a 5% polyacrylamide, 8 M urea gel[13] The regions corresponding to bands II and III (as visible in *a*) are indicated, the lane between bands II and III which has the least amount of radioactivity and the two lanes corresponding to the two fractions above band II were not assigned to either band T, unfractionated reaction The structure of the different RNAs, as deduced from their electrophoretic mobility on this and other gels of different polyacrylamide concentrations, is shown at the side

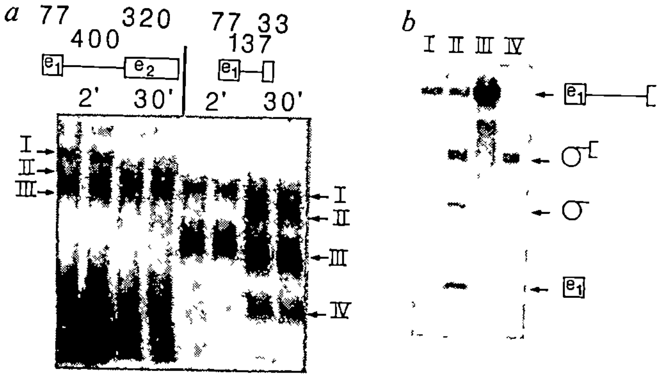


Fig. 4 Distribution of pre-mRNA, intermediates and products among splicing complexes formed with a short substrate. *a*, Preparative, non-denaturing gel showing the analysis of splicing complexes formed in the presence of a full-length (rp51A, 800 nt long) or short (rp51A $\delta 2/DdeI$, 247 nt long) RNA substrate after 2-min or 20-min incubations in a splicing reaction (each reaction was loaded on two lanes). An autoradiogram obtained after electroblotting the gel onto a DEAE membrane[13] is shown. The structure of the transcripts is shown above the lanes. *b*, The RNAs were extracted from each band and analysed on a 5% polyacrylamide, 8 M urea gel[13]. The structure of the different RNAs as determined by their electrophoretic mobility is shown on the right.

yeast spliceosomes[13]. Comparison of our patterns of immunoprecipitated snRNAs with those in a recently reported study[23] suggests that these three species are those known as snR14 (160 nt), $snR7_S$ (185 nt) and $snR7_L$ (215 nt). The latter two are transcripts from the SNR7 gene and vary in the lengths of their 3' ends[23].

An additional, unusually large, trimethyl-capped species (1,100–1,200 nt) also appeared to be associated with complexes formed on a polyadenylated substrate, but not with spliceosomes purified by gel electrophoresis, probably because of a high background level in that region of the gel[13]. In the presence of carrier RNA the splicing complexes migrate in a region of the gel showing less background, allowing us to confirm that this RNA is also part of the spliceosome. Based on its molecular weight, it is almost certain that this large RNA corresponds to a recently discovered essential RNA species, lsr1 (large spliceosomal RNA1), the product of the LSR1 gene, which has extensive sequence homology to the mammalian U2 snRNA[24]. Comparison of gel patterns suggests that this species is identical to snR20[23]. Although these comparisons are not yet definitive, we will refer to these RNAs as LSR1, $snR7_L$, $snR7_S$ and snR14.

We analysed the distribution of these RNAs among the different splicing complexes, comparing their level in each complex with the background level in the absence of *in vitro* assembled complexes (see Fig. 5 and ref. 13). By this criterion, each complex contains a specific complement of trimethyl-capped RNA species. lsr1 is clearly present in all three complexes, band III appears to contain only this species, as determined by this assay. Band I also contains the three other RNAs previously shown to be associated with spliceosomes, ($snR7_S$, $snR7_L$ and snR14). Band II contains lsr1, $snR7_S$ and $snR7_L$ but is missing snR14. Thus, we can account for all the species described previously and the three complexes show different snRNA compositions (Fig. 6).

Pathway of complex formation

The addition of high concentrations of competitor RNA before electrophoresis greatly improves the resolution of splicing complexes over our previously reported method[13] and allows a more detailed description than has been possible with the more conventional technique of sedimentation analysis[14–17]. The competitor RNA presumably removes non-specific or relatively loosely-associated factors, reducing the size and increasing the electrophoretic mobility of the splicing complexes. The relative

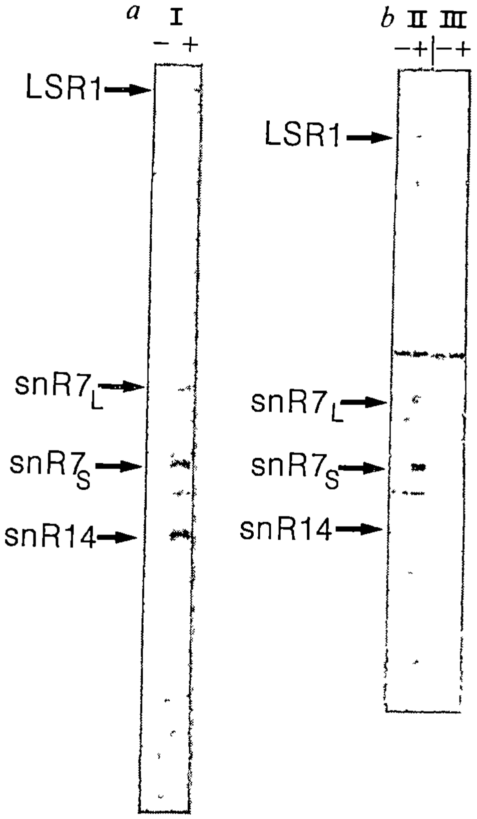


Fig. 5 Trimethyl-capped RNAs in the three splicing complexes. snRNAs present in bands I, II and III. Preparative reactions in the presence of the short substrate (rp51A $\delta 2/DdeI$) were incubated for 2 and 30 min at 25 °C and run on gels similar to those shown in Figs 3 and 4, after incubation for 10 min on ice in the presence of 1 μg μl^{-1} of periodate-treated *Xenopus laevis* RNA (see below). The gels were then electroblotted onto DEAE membranes and the RNAs present in band I (panel *a*) or bands II and III (panel *b*) were eluted from each lane, ethanol-precipitated in the presence of 5 μg of salmon sperm DNA (see below), end-labelled with pCp and RNA ligase (as described[31] with the modifications described in ref. 13), immunoprecipitated with anti-trimethyl-cap rabbit antibody (as described[32], with the modifications described in ref. 13) and analysed on a 7% polyacrylamide 8 M urea gel. To ensure that the capped RNA species are specifically associated with *in vitro* assembled complexes, a reaction containing no exogenously-added substrate was processed in parallel[13]. The lanes containing complexes I, II and III are indicated. + and − denote the presence or absence of exogenously-added substrate in the splicing reaction. The sizes of the various RNAs were determined (on this and other gels) by comparison with 5S RNA, 5.8S RNA, the rp51A $\delta 2/DdeI$ transcript (247 nt), the rp51A transcript (800 nt) and 18S RNA. Two modifications were introduced in these experiments to ensure that the snRNAs observed were present in splicing complexes and not artefactual. First, *Xenopus laevis* RNA that had been oxidized with periodate (to make it unavailable for end-labelling with pCp and RNA ligase[31]) was used as competitor instead of total yeast RNA. The protocol for periodate treatment was provided by M. Green. Briefly, 200–500 μg of total *Xenopus laevis* RNA were resuspended in 1 ml of 0.3 M NaOAc *p*H 5.3, 1 ml of (4 mM $NaIO_4$, 10 mM NaOAc *p*H 5.3) was added and the mixture was incubated in the dark at 4 °C for 1 h, then at room temperature for 1 h. The reaction was quenched by adding 0.2 ml of propylene glycol (1,2-propanediol) and incubated at room temperature for an additional 10 min. The RNA was then phenol-extracted, ethanol-precipitated and washed extensively. Second, instead of tRNA as carrier in the ethanol precipitations after elution from the DEAE membrane, we used salmon sperm DNA which had been previously freed of any contaminating RNA by overnight incubation in 0.3M NaOH and ethanol precipitation in the presence of 2.5 M NH_4Ac, because we found that salmon sperm DNA was not (or only very poorly) labelled in a standard pCp end-labelling reaction.

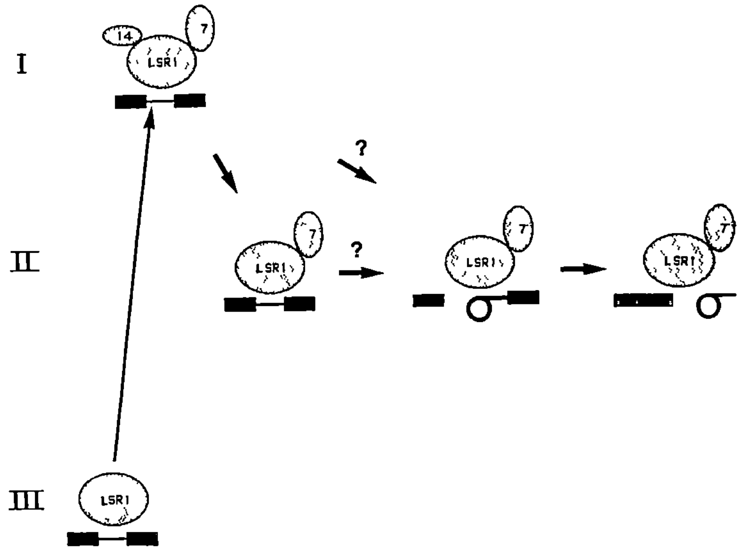


Fig 6 Schematic summary of the data corresponding to the complexes formed on the pre-mRNA and those containing the products of the first and second step Each band (I, II and III) is indicated with the complex(es) it contains The composition in snRNAs is based on the assumption that a given band is homogeneous, which may not be the case The arrows indicate transformation of one complex into another, which might involve other undetected intermediate complexes The question marks reflect the fact that it is not clear which complex is the site of the first step

mobility of the three major complexes is inversely related to the number of snRNAs found associated with each of them It is not clear whether this is a causal relationship as the snRNA characterization includes a number of uncertainties There may be additional snRNAs, undetected in this and our previous analysis (see discussion on this point in ref 13) and differences in protein composition and/or conformation between the three complexes

We suggest that the three complexes form in an ordered manner III→I→II To demonstrate such a pathway rigorously, one would have to show that it is possible to 'chase' pre-mRNA through the different complexes This has not yet been possible but three lines of evidence, taken together, suggest that this interpretation is correct

First, the kinetics of appearance of the various complexes show clearly that bands I and III appear early in the reaction, preceding any covalent modification of the substrate Band II appears subsequently, as the products of the first step, 5′ exon and lariat intermediate, first become detectable Although the precise kinetics depend on the extract used, temperature and other details of the individual experiment, the two early bands (III and I) always appear together under normal conditions In addition, both bands usually decrease substantially as band II intensifies (Fig 2*a*) In several experiments in which the substrate pre-mRNA was saturating, band III accumulated to high levels and did not decrease (Figs 3*a* and 4a), even after long incubations (data not shown) In contrast, band I is only easily detectable at early times of incubation Second, an examination of assembly under non-optimal conditions indicates that only band III is formed This is true with a substrate containing a single point mutation at the branchpoint (Fig 2*b*) as well as with defective extracts obtained, for example, by diluting a normal extract (data not shown) The III→I→II sequence explain the data most simply but a model where either III or I can lead to II is also possible Analysis of the snRNAs associated with each complex suggests that III is the initial complex as it contains lsr1 but not snR7 (S or L) or snR14 Band I is then formed with the binding of these two other snRNAs, and band II follows with the loss of snR14

Sites of the two steps of splicing

The complex in which the first step occurs is expected to contain pre-mRNA and the products of this reaction Both bands I and II are candidates for this active complex since we cannot resolve the two complexes clearly, they both appear to contain the two types of molecules (The products of the first step are much more abundant in band II than band I, however) If band I is the site of the first step, 5′ cleavage and lariat formation are accompanied or rapidly followed by the loss of snR14 and the transition to band II, which then more slowly undergoes the second step If band II is the site of the first step, the presence of pre-mRNA in band II may be due to the fact that, at least *in vitro*, this active complex can also be formed (for example, band I can lose snR14) without undergoing the first step We therefore cannot determine whether the loss of snR14 and the concomitant band I→band II transition occur before or after the first step of the reaction

The site of the second step of the reaction, 3′ cleavage and exon ligation, can be assigned more easily to band II The products of the first step (intron-exon 2 lariat and exon 1), which are the substrates for the second step, are found in highest concentration in band II The distribution of the products of the second step is also compatible with this view In contrast to the lariat intron, almost none of the exon 1-exon 2 RNA is found associated with the large splicing complexes[13,14,16,17] In Fig 3 a low level of exon 1-exon 2 RNA is detected and coincides with band II A similarly low level of excised intron is also present in the band II fractions, larger amounts of this molecule are in faster migrating complexes which may be derived from band II simply by loss of the exon 1-exon 2 RNA

This description does not relate the complexes observed to the *in vivo* complexes, or even to the complexes formed in *in vitro* splicing, from which bands I-III are derived The treatment involved is particularly severe, including high salt, the addition of EDTA and, notably, incubation with non-specific competitor RNA (see Fig 1) and can be expected to disrupt all but the strongest interactions Thus the RNP gels show a 'core' splicing complex This may affect the interpretation of the proposed assembly pathway As an extreme possibility, there may exist a preformed spliceosome containing all the different snRNPs that will bind to the pre-mRNA, the interactions between the different snRNPs and/or between certain snRNPs and a given part of the pre-mRNA substrate might then become stronger or weaker as the splicing reaction proceeds The different complexes observed here would then reflect different levels of interaction rather than the step by step association or disassociation of snRNPs with a complex containing the pre-mRNA substrate Further studies may lead to an understanding of the interactions which produce stable complexes and the active splicing complexes, and the identification of the factors which recognize the conserved sequence elements in yeast introns

This study was supported by grant GM23549 from the NIH to M R C W P was supported by a Dretzin predoctoral fellowship and B C R by a post-doctoral fellowship from the NIH We thank A Jacquier and M Green for helpful discussions, D Torrey for technical assistance, R Luhrmann for anti-TMG antibodies, and T Tishman for typing the manuscript

1 Ruskin, B , Krainer, A R., Maniatis, T & Green, M R. *Cell* **38**, 317-331 (1984)
2 Padgett, R. A , Konarska, M , Grabowski, P J , Hardy, S F & Sharp, P A *Science* **225**, 898-903 (1984)
3 Konarska, M M , Grabowski, P J , Padgett, R A & Sharp, P A *Nature* **313**, 552-557 (1985)
4 Lin, R-J , Newman, A J , Cheng, S C & Abelson, J *J biol Chem* **260**, 14780-14792 (1985)
5 Newman, A J , Lin, R J , Cheng, S C & Abelson, J *Cell* **42**, 335-344 (1985)
6 Domdey, H *et al. Cell* **39**, 611-621 (1984)
7 Rodriguez, J R., Pikielny, C W & Rosbash, M *Cell* **39**, 603-610 (1984)
8 Zeitlin, S & Efstratiadis, A *Cell* **39**, 589-602 (1984)
9 Padgett, R A , Grabowski, P J , Konarska, M M , Seiler, S & Sharp, P A *A Rev Biochem.* **55**, 1119-1150 (1986)
10 Green, M R. *A Rev Genetics* **20**, 671-708 (1986)
11 Rymond, B C & Rosbash, M *Nature* **317**, 735-737 (1985)
12 Ruskin, B , Pikielny, C W , Rosbash, M & Green, M R. *Proc natn. Acad Sci USA* **83**, 2022-2026 (1986)
13 Pikielny, C W & Rosbash, M *Cell* **45**, 869-877 (1986)
14 Brody, E & Abelson, J *Science* **228**, 963-967 (1985)
15 Grabowski, P J , Seiler, S R. & Sharp, P A *Cell* **42**, 345-353 (1985)
16 Frendewey, D & Keller, W *Cell* **42**, 355-367 (1985)
17 Bindereif, A & Green, M R. *Molec. cell Biol* **6**, 2582-2592 (1986)
18 Vijayraghavan, U *et al. EMBO J* **5**, 1683-1695 (1986)
19 Reddy, R. & Busch, H in *Progress in Nucleic Acids Research and Molecular Biology* **30**, (Cohn, W E , ed) 127-162 (Academic, New York, 1983)
20 Cheng, S -C & Abelson, J *Proc natn Acad Sci USA* **83**, 2387-2391 (1986)
21 Jacquier, A & Rosbash, M *Proc natn. Acad Sci USA* **83**, 5835-5839 (1986)
22 Pikielny, C W , Teem, J L & Rosbash, M *Cell* **34**, 395-403 (1983)
23 Riedel, N , Wise, J A , Swerdlow, H , Mak, A & Guthrie, C *Proc natn. Acad Sci USA* **83**, 8097-8101 (1986)
24 Ares, M Jr *Cell* **47**, 49-59 (1986)
25 Mount, S M , Pettersson, I , Hinterberger, M , Karmas, A & Steitz, J A *Cell* **33**, 509-518 (1983)
26 Black, D L , Chabot, B & Steitz, J A *Cell* **42**, 737-750 (1985)
27 Chabot, B , Black, D L , LeMaster, D M & Steitz, J A *Science* **230**, 1344-1349 (1985)
28 Ruskin, B & Green, M R. *Cell* **43**, 131-142 (1985)
29 Melton, D A *et al. Nucleic Acids Res* **12**, 7035-7056 (1984)
30 Pikielny, C W & Rosbach M *Cell* **41**, 119-126 (1985)
31 England, T E , Bruce, A G & Uhlenbeck, O C *Meth. Enzym.* **65**, 65-74 (1980)
32 Bringmann, P *et al. J biol Chem* **258**, 2745-2747 (1983)

LETTERS TO NATURE

An accretion event in the Seyfert galaxy NGC 5548

B. M. Peterson & G. J. Ferland

Department of Astronomy, The Ohio State University, 174 West 18th Avenue, Columbus, Ohio 43210, USA

Accretion on to a massive collapsed central object is now commonly thought to power active galactic nuclei (AGN)[1–3]. If this is indeed so, then changes in luminosity of the system should correspond to changes in accretion rate, or, equivalently, the mass flux at regions a few Schwarzschild radii from the putative black hole. The observational consequences of such a change in accretion rate may not be pronounced, as gas near the central source of radiation is expected to be very highly ionized and difficult to observe outside the X-ray band. Here we draw attention to a remarkable change which occurred in the spectrum of the low-luminosity, optically variable, Seyfert galaxy NGC 5548. An increase in optical luminosity was accompanied by the appearance of a strong, abnormally broad He II 4,686-Å emission line. The appearance of this line indicates the presence of a component of highly ionized gas that did not contribute to the optical spectrum before the increase in luminosity. The greater line width suggests that the gas producing the He II emission is much closer to the central object than the broad region (BLR). This observation may be the first direct evidence of the accretion onto a central object.

We have been observing the low-luminosity Seyfert 1 galaxy NGC 5548 for some time[4] Optical spectra have been obtained at regular intervals A major goal of this work was to determine time lags between continuum activity and the response of the emission-line region The time delay, which establishes the separation between the central engine and the BLR, is found to be around 31 days, corresponding to $r \simeq 7\ 9\times10^{16}$ cm (ref 5) If the BLR gas is bound in the potential well of the central object then the corresponding virial mass is $\sim 3\times10^{7}\ M_{\odot}$

During the course of these observations, the optical continuum increased by around 66% (see Fig 1, in which the flux density of the continuum near 5,500 Å is shown, after removal of the underlying galaxy continuum) Details of the observational procedure and reduction techniques are discussed elsewhere[4] This increase in visual brightness was accompanied by a similar increase in the extreme ultraviolet (EUV, that is $\lambda <$ 912 Å) because the flux in the Balmer lines also increased (when the 31 day time-lag is taken into account) The fact that the increase in brightness in the continuum and Balmer lines increased together suggests that the luminosity of the EUV and optical continua are well correlated

The surprising feature of these observations is the unexpectedly large increase in the strength of the He I 4,686-Å line (see Fig 1) This phenomenon has not been reported before, although it is only within the last few years that Seyfert galaxies have been spectroscopically monitored with sufficient frequency to make such detection possible Whereas the continuum increased by a factor of ~1 7, the He II line increased more than tenfold The most likely explanation for this change was the addition of a new gaseous component He II lines are formed by recombination in active nuclei[6] the luminosity of a recombination line will remain constant if the gas is fully ionized and the mass and density of the emitting gas remain constant In a radiation-bounded geometry, one in which the majority of the ionizing radiation striking a BLR cloud is absorbed by the emitting gas, the equivalent width of an emission line should remain constant, and the He II line should increase in phase with the continuum Our observations rule out both these simple pictures

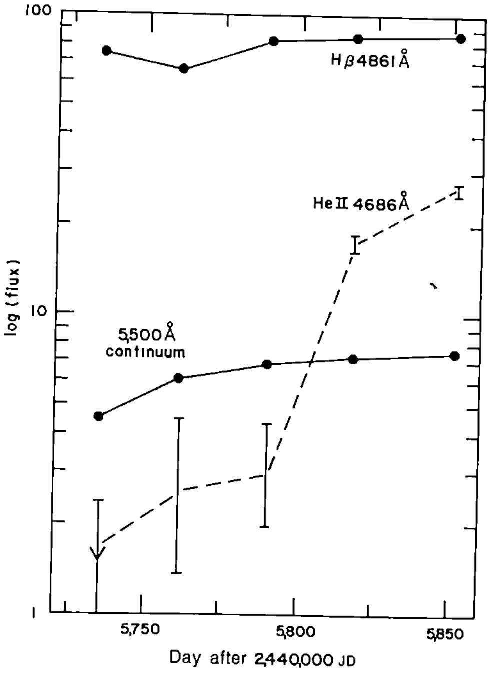


Fig 1 The flux in Hβ, He II 4,686 Å and the 5,500 Å continuum, over the course of our observations Line fluxes are in units of 10^{-14} erg s^{-1} cm^{-2} and continuum fluxes are in units of 10^{-15} erg s^{-1} cm^{-2} Å^{-1}

It is possible to argue that the dramatic increase at 4,686 Å is caused by a flattening of the continuum, thereby increasing the number of helium-ionizing photons relative to the optical continuum We reject such an explanation because our data show no evidence for continuum shape change at lower energies, based on ultraviolet and optical observations Moreover, the

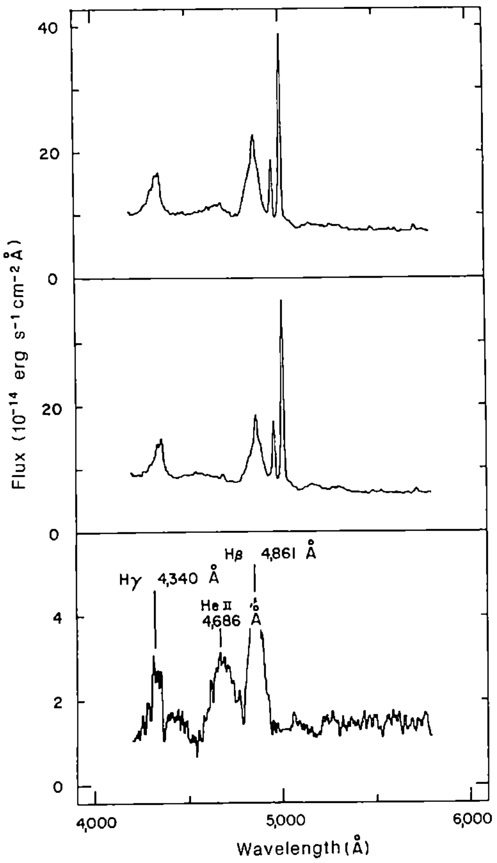


Fig 2 The top two panels show the optical spectrum near the Hβ and He II 4,686 Å lines during the high and low states The bottom panel is a difference spectrum, with the strongest emission lines marked The high-state spectrum was obtained on 1-2 June 1984 UT (2,445,852-3 JD) and the low-state spectrum was obtained on 2-3 March 1984 UT (2,445,761-2 JD)

Balmer lines and the He I 5,876-Å line, which are sensitive to the radiation field in the range 1-4 Ryd, did not increase by as large a fraction as the optical continuum if the increase at 4,686 Å is attributable to an increase in the density of ionizing photons, the increase in flux would have to be confined to energies in excess of 4 Ryd We therefore believe that the increase in the I(4,686 Å)/I(Hβ) ratio was caused by the appearance of a new component of gas

The He II 4,686-Å line profile confirms that a new component of gas is present, that this gas is not associated with the classical BLR, and that the gas is probably closer to the central object that the BLR Figure 2 shows a difference spectrum, obtained by subtracting a low-state spectrum (when He II emission is weak) from one obtained in the high state, when the anomalous He II emission was present The increased Hβ line is apparent, along with Hγ near the blueward limit of the spectrum The Hβ profile is similar to that seen in low-state spectra, and is an indication of the width of broad emission lines in this object Surprisingly, the new He II emission line has a width ~75% greater than the neighbouring Hβ Although there is some tendency for He II lines to be slightly broader than hydrogen lines[7], the difference seen here is much greater than has ever been seen before The greatly enhanced He II emission can be attributed to gas moving with much greater velocity than BLR clouds If the gas is confined within the potential well of the central object, we can infer that the He II-emitting material lies at roughly one-third of the BLR radius from the central object, or at roughly 3×10^{16} cm We thus conclude that the increase in luminosity of the central object in NGC 5548 was accompanied by the appearance of a new component of gas much closer to the central object than previously detected The coincidence further suggests that the newly present very-broad-line-region (VBLR) gas may be associated with the increase in accretion rate that must occur in the standard picture

At first sight, it seems surprising that the VBLR is detected only in He II The lack of a correspondingly spectacular change in Hβ, or He I 5,876 Å is convincing evidence that the gas is fully ionized, namely, He is doubly ionized throughout the VBLR If so, then the I(4,686 Å)/I(Hβ) intensity ratio would be near its case B value of ~1 for a solar composition Strong VBLR Hβ emission is not expected Although Hβ appears much sharper than He II 4,686 Å, tests show that a fairly substantial component of very broad Hβ could exist in the observed profile Line profile synthesis shows that the I(4,686 Å)/I(Hβ) ratio in the VBLR lies between 1 and 2 If both lines radiate with case B emissivities then this ratio sets the limits $0\,1 \leq \text{He/H} \leq 0\,2$, that is, the material does not seem to be highly processed

A simple argument suggests that the new component of gas nearly covers the continuum source completely For conditions in the BLR, where the gas is optically thick to ionizing radiation, the equivalent width of Ly-α is related to the mean optical-X-ray continuum slope by $W(\text{Ly-}\alpha) = (1{,}216/\alpha)(912/1{,}215)^{\alpha}(\Omega/4\pi)$, where α is the spectral index and $\Omega/4\pi$ is the covering factor of absorbing gas This assumes that each ionizing photon eventually produces one Ly-α The spectral index is[8] $\alpha_{\text{ox}} \simeq 1\,1$, corresponding to a BLR covering factor of ~15% This value is fairly typical of AGN The He II I(4,686 Å)/I(Hβ) ratio before the increase in luminosity was also fairly typical of active nuclei (see Fig 1) If we take the pre-increase 4,686-Å flux as typical of gas with a 15% covering factor, then the roughly sevenfold increase in equivalent width at 4,686 Å corresponds to an increase in the covering factor to 50-100% The increased equivalent width further suggests that the VBLR gas absorbs the majority of the He$^+$-ionizing radiation ($\lambda \leq 228$ Å) This fact, together with the large covering factor, suggests that it is this ephemeral VBLR that may account for the variable low-energy X-ray absorption characteristic of low-luminosity Seyfert galaxies[9] The fact that the VBLR does not contribute He I lines (He I lines are present in our spectra, but have widths similar to the Balmer lines) shows that the VBLR is not, however, optically thick to He$^+$-ionizing radiation

Simple calculations allow estimation of the column density and mass of the VBLR gas For an assumed distance of 70 Mpc the 4,686-Å flux at maximum light is $\sim1\,0\times10^{41}$ erg s^{-1} This corresponds to an emission measure of $N^2V = 8\times10^{66}$ cm^{-3}, assuming a temperature of 10^5 K and He/H = 0 1, where N is the hydrogen density and V the emitting volume There is no reliable density indicator in the VBLR spectrum that we observe, but if we assume a density of 10^{10} cm^{-3} (a typical BLR density) and full coverage at a distance of 3×10^{16} cm then the VBLR gas has a column density of $\sim7\times10^{22}$ cm^{-2}, typical of X-ray absorbing gas The geometry of this gas is unclear because unpublished photoionization calculations show that such clouds would be unstable to disruption by radiation pressure The emission measure also corresponds to a mass of 0 8 $M_{\odot}$, assuming a solar composition An obvious possibility is that the rapid appearance of the VBLR was the result of the disruption of a star Such conjectures are beyond the scope of this report, but a possibility is that a star was disrupted in the inner regions of the system when the optical continuum began to rise, but that the newly added gas only became spectroscopically active in the optical region when debris was ejected to nearly the BLR radius

We have shown that a new emission-line region, not previously detected in the optical spectrum, appeared when the non-thermal continuum in NGC 5548 increased in luminosity If current ideas about the energy source in AGN are correct, then such an increase in luminosity can only be caused by an increase in the accretion rate in the innermost regions of the nucleus We believe that we have actually observed such an accretion event in progress

Received 14 July, accepted 1 September 1986

1 Lynden Bell, D *Nature* **223**, 690-694 (1969)
2 Rees, M J *Physica Scr* **17**, 193-200 (1978)
3 Begelman, M C in *Astrophysics of Active Galaxies and Quasi Stellar Objects* (ed Miller, J) 711-752 (University Science Books, Mill Valley, California, 1985)
4 Peterson, B M *Astrophys J* **312** (in the press)
5 Peterson, B M & Gaskell, C M *Astr J* **91**, 552-556 (1986)
6 MacAlpine, G M , Davidson, K., Gull, T R & Wu, C C *Astrophys J* **294**, 147-152 (1985)
7 Osterbrock, D E & Shuder, J *Astrophys J Suppl* **49**, 149-174 (1982)
8 Mushotzky, R. F, Marshall, F, Boldt, E, Holt, S & Serlemitsos, P *Astrophys J* **235**, 377-385 (1980)
9 Reichert, G A, Mushotzky, R F, Petre, R & Holt, S S *Astrophys J* **296**, 69-89 (1985)

Large-amplitude semidiurnal temperature variations in the polar mesopause: evidence of a pseudotide

R. L. Walterscheid*, G. G. Sivjee*, G. Schubert* & R. M. Hamwey†

* Space Sciences Laboratory, The Aerospace Corporation, POB 92957, Los Angeles, California 90009, USA
† Department of Physics and Astronomy, Dartmouth College, Hanover, New Hampshire 03755, USA

Current atmospheric tidal models[1,2] predict an amplitude of ≲2 K for the solar semidiurnal temperature variations in the polar mesopause (~85 km altitude) during the winter solstice period. In contrast, here we report an amplitude of about 13 K for the 12-h variations in the mesopause temperature at Longyearbyen, Spitsbergen (78.2° N) on UT day 359 in 1984. The kinetic temperature of the mesopause was deduced from the thermalized rotational distribution of the mesopause airglow OH(8,3) vibrational–rotational band emission in the near-infrared spectral region. A semidiurnal 'pseudotide' is proposed as the source of the observed large-amplitude 12-h temperature variation. This pseudotide is generated by gravity-wave momentum fluxes modulated by the true semidiurnal tide[3]. The semidiurnal pseudotide is very efficiently excited at very high latitudes. Based on conservative values of gravity-wave amplitudes, we estimate a temperature amplitude of ~10 K.

Certain exothermic chemical reactions among mesopause constituents, particularly the interaction of O_3 with H, produce OH in vibrationally and rotationally excited states The radiative lifetime of these radicals is ~1 ms, whereas the collision frequency in the mesopause is $\sim 10^4 \, s^{-1}$ Consequently the OH radicals undergo ~10 collisions, on average, before shedding their excess energy, and their distribution among the rotational states is thermalized to the ambient kinetic temperature Measurements of the relative intensities of the airglow OH rotational-line emissions can therefore provide a measure of the mesopause kinetic temperature

Spectroscopic observations of the OH rotational lines in the P_1 and P_2 branches of the (8,3) and the (6,2) vibrational–rotational bands were made in Longyearbyen, Spitsbergen, in December 1984 The high latitude of the observatory (78 2° N) permitted continuous 24-h observations of the wintertime polar airglow On UT day 359 in 1984 the spectroscopic observations around 7,300 and 8,500 Å showed no auroral band emissions or O II 7,320-Å line emission, and only a weak O I 8,446-Å line emission The absence of a significant level of any auroral optical emission indicates that there were no particle precipitation events during this period energetic enough to have had any effect on the thermodynamics of the mesopause The local solar shadow height above the Longyearbyen observatory was >150 km over the entire observing period

We analysed the relative intensities of the spectrally resolved OH rotational lines, using Mies's[4] transition probabilities and Kvifte's[5] as well as Coxon and Foster's[6] rotational-term values and emission wavelengths, to derive the mesopause temperature, T The average temperature, $\bar{T}$, of the mesopause over the 24-h period on UT day 359 in 1984 was 219±2 K For comparison with other reports of mesopause temperature variations (albeit of shorter period, see, for example, Krassovsky[7]), we have analysed the normalized changes in temperature, that is, $(T-\bar{T})/\bar{T}$ Figure 1 shows the time history of this quantity for the 24-h period on UT day 359 in 1984, and also the 12-h component of the temporal spectrum calculated by Fourier analysis of the temperature-data set The major mode of the wave-like structure in the time history of $(T-\bar{T})/\bar{T}$ is clearly the semi-diurnal wave with an amplitude of 0 0574, which corresponds to $T-\bar{T} \simeq 13$ K

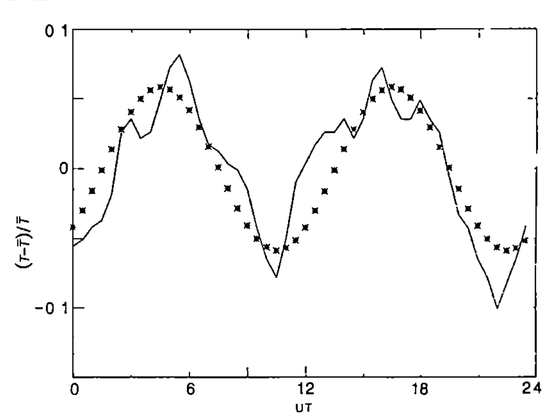


Fig 1 Time history of normalized changes in mesopause temperature, that is, $(T-\bar{T})/\bar{T}$, for a 24-h period on UT day 359 in 1984 (continuous curve) The contribution attributable to the 12-h Fourier component (*) is also shown

Forbes's[2] model calculations of the semi-diurnal components of the solar and lunar atmospheric tides predict amplitudes of 2 0 and 0 4 K at 87 km height in the polar region (78° N) during the December solstice period Our calculations, using the model described in ref 1, predict even smaller amplitudes (<1 K) This model includes a large number (8) of the conventional zonal wavenumber 2 semidiurnal modes Clearly, the normal tidal oscillations of the atmosphere do not account for the order-of-magnitude larger amplitude of the observed semidiurnal temperature oscillations The observed temperature oscillations also cannot be due to very-high-order solar-driven modes, which potentially have significant amplitudes at high latitudes, because these modes are only weakly excited by global solar forcing, and local solar forcing is lacking in winter The observed temperature oscillations might be due to the viscous distortion of the principal solar-driven modes This distortion can broaden the latitude region having significant amplitude However, such strong broadening has not been seen at emission altitudes in a model that includes this effect[2]

The occurrence of semidiurnal oscillations in polar-latitude mesopause temperatures is not an unusual event Myrabø[8] co-added 19 days of airglow data from Spitsbergen for the period 9-27 December 1982 and found a semidiurnal temperature amplitude of ~3 K, but no diurnal variation Viereck *et al* (R A Viereck, C H Deehr, G G Sivjee and G J Romick, manuscript to be submitted to *Planet Space Sci*, hereafter referred to as VDSR) report that semidiurnal oscillations are seen in more than half the OH airglow data obtained from Spitsbergen during the winters of 1984-85 and 1985-86 In some of the data the semidiurnal oscillations are quite large, the co-addition of five consecutive days of data gave a predominantly semidiurnal variation with a peak-to-peak amplitude of ~16 K Forbes[2], on the other hand, predicts a predominantly diurnal variation at the latitude of Spitsbergen VDSR do report the presence of diurnal variations in their inferred mesopause temperatures

The above considerations suggest the possibility of a high-latitude source for the oscillations such as Joule and particle heating in the lower thermosphere However, conventional semidiurnal modes are not good candidates for exhibiting the effects of high-latitude forcing because, even though extremely high-order modes can potentially have appreciable amplitude at high latitude, they are strongly absorbed by dissipative processes in the lower thermosphere and, moreover, are only weakly excited by high-latitude heating Because these modes are not confined to high latitude, but are global in extent, high-latitude heating does not project well onto these modes In fact, the totality of conventional semidiurnal modes (denoted gravitational modes by Flattery[9]) are not able to resolve very-high-latitude thermal forcing Gravitational modes alone are not complete, and another set of modes called rotational modes[9] are required when $\lambda = \omega/2\Omega < 1$, where ω is the wave angular frequency and Ω is the Earth's angular speed For solar semidiurnal tides $\lambda = 0\ 9973$ Rotational modes with $\lambda \simeq 1$ are trapped very close to the pole[10] The main difficulty with invoking rotational modes as the source of the observed temperature variation is that these modes are evanescent and have very short vertical scales ($\lesssim$300 m), they would therefore require very strong local forcing, which does not exist at mesopause altitudes during polar night These modes might be viable if strong mean wind effects altered the effective rotation rate and if the OH airglow emissions were coming from altitudes much higher than now believed Also significant in the data of Fig 1 is the lack of any substantial power at the frequency of the diurnal tide If the semidiurnal tide were responding to high-latitude forcing then the diurnal tide ought to be even more prominent because certain diurnal modes respond more efficiently to high-latitude sources than do the semidiurnal modes

It is possible that the semidiurnal variation reflects near-inertial oscillations, because the inertial period at Spitsbergen is nearly semidiurnal (12 3 h) However, it seems unlikely that free inertial waves would show the phase coherence seen in Fig 1 over two complete cycles Moreover, inertial motions are mainly horizontal and are associated with small temperature fluctuations Also, near-inertial waves have very short vertical wavelengths, but the time series for emission intensity for two different OH bands ((8,3) and (6,2)) show semidiurnal variations that are essentially in phase, although the emission heights for the two bands are separated by a few kilometres

We propose that the observed semidiurnal temperature oscillation is a pseudotide driven by gravity-wave-mean-flow interactions modulated by the solar-driven tide The pseudotide is generated as follows the rate at which gravity waves force the mean flow increases as $|\bar{u}-c|$ decreases, where $\bar{u}$ is the component of the large-scale flow in the direction of wave propagation, and c is the phase velocity of the wave, both quantities are positive eastward For convenience, consider zonally propagating two-dimensional waves with $\bar{u}$ the zonal component of the semidiurnal tide Waves with $c > 0$ ($c < 0$) impart an eastward (westward) acceleration to $\bar{u}$ During the eastward (westward) phase of the semidiurnal oscillation eastward (westward) propagating waves are preferentially absorbed The net effect is to induce a semidiurnal contribution to $\bar{u}$ which is non-tidal This mechanism is discussed in more detail in ref 3

In the simple case of zonally propagating two-dimensional gravity waves the forcing of the mean flow can be expressed as $\partial^2\bar{u}/\partial t^2 + f^2\bar{u} = (\partial/\partial t)\bar{G}$, where $\bar{G} = -(1/\rho_0)(\partial/\partial z)\rho_0\overline{u'w'}$, $f = 2\Omega \sin\phi$ is the Coriolis parameter (ϕ is latitude), ρ_0 is a reference density profile, z is height, and u' and w' are respectively the zonal and vertical gravity-wave velocities The quantity $\bar{G}$ is the forcing of $\bar{u}$ due to the wave-momentum flux convergence An overbar refers to a spatial average taken over a distance equal to the wavelength of the longest gravity waves making significant contributions to the momentum flux When $\bar{G}$ is periodic with semidiurnal frequency $\omega \simeq 2\Omega$, $\hat{u} = -[i\omega/(\omega^2 - f^2)]\hat{G} \simeq -(1/\cos^2\phi)(1/\omega)\hat{G}$, where $\hat{G}$, $\hat{u}$ are the Fourier transforms of $\bar{G}$, $\bar{u}$ At the latitude of Spitsbergen the amplification factor $\alpha = |\omega^2/(f^2 - \omega^2)|$ ($= 1/\cos^2\phi$, for $\omega = 2\Omega$) is 23 for the semidiurnal frequency By contrast, $\alpha = 0\ 35$ for the diurnal frequency At 60° N, $\alpha = 4$ for the semidiurnal frequency Thus strong amplification in polar regions is limited to the semidiurnal frequency, and to very high latitude A semidiurnal oscillation with an amplitude of $\sim$10 m s^{-1} could be induced near mesopause heights for conservative gravity-wave amplitudes[3] and $\alpha \sim 2$

Thus it seems possible for a gravity-wave semidiurnal tide interaction at very high latitudes to force a large-amplitude pseudotide The plausibility of the interaction is enhanced by the relatively large semidiurnal tidal-wind speeds Semidiurnal tidal-wind velocities are not proportionately as small as temperature oscillations at high latitudes because the latitudinal structure functions for the wind components decay toward the poles much more slowly than do the temperature functions (that is, Hough functions) Our tidal model predicts semidiurnal tidal velocities of $\sim$20 m s^{-1} near 90 km at 75° N during the winter This estimate, together with the results presented in ref 3 for a gravity-wave phase speed of $\sim$20 m s^{-1}, suggest pseudotide amplitudes of $\sim$50 m s^{-1} at Spitsbergen, provided that momentum fluxes are comparable In the calculations gravity-wave amplitudes were 10 m s^{-1} at 80 km, compared with observed r m s amplitudes of $\sim$20–40 m s^{-1} (refs 11, 12) Amplitudes at very high latitudes have not been measured, but the latitude dependence is apparently not strong[11-14] Wintertime observations at Heiss Island (80 4° N) show semidiurnal wind amplitudes of $\sim$10 m s^{-1} (ref 15) These values are averages over the meteor region ($\sim$80–100 km) and are monthly averages based on three winters Because the dominant tidal modes at high latitude are apt to have short vertical wavelengths ($\sim$20 km) and because the induced tide is apt to be rather variable in time (reflecting variable gravity-wave fluxes), predictions of large semidiurnal velocities are consistent with the observations

To estimate the induced oscillation of temperature we must take into account divergent motions that result when mean quantities vary in latitude The temperature change, $\bar{T}$, resulting from adiabatic heating and cooling can be estimated as follows From the continuity equation $\bar{w} \sim (M/L)\bar{u}$, where L is the meridional length scale and M is the vertical length scale In obtaining this equation we have used $\bar{u} \sim \bar{v}$, where $\bar{v}$ is the meridional wind component From the first law of thermodynamics $\bar{T} \sim (1/\omega)\Gamma_a[(M/L)\bar{u}]$, where the quantity in square brackets is $\bar{w}$ and where Γ_a is the adiabatic lapse rate ($= g/c_p$, where g is the acceleration of gravity and c_p is the specific heat of air at constant pressure) With $\Gamma_a = 10^{-2}$ K m^{-1}, $M \sim 20 \times 10^3$ m$/2\pi \sim 3 \times 10^3$ m, $L \sim 10^6$ m (L is the half width of the region over which α is large and the driving tide is sensible, the tidal amplitudes vanish at the pole) and $\bar{u} \sim 50$ m s^{-1}, we obtain $\bar{T} \sim 10$ K This estimate agrees well with the observed amplitude of 13 K It is also consistent with the results of Myrabø[8] and VDSR mentioned above These results were based on averages over a number of days because the day-to-day variability of the semidiurnal variation in temperature can be quite large This large temperature variability perhaps reflects the unsteadiness of the gravity wave fluxes[3]

We note finally that some authors have observed a connection between stratospheric warmings and changes in the amplitude of the semidiurnal tide (ref 16 and VDSR) The observations presented here were made one day before the onset of a sudden warming was observed in stratospheric data[17] and several days before the characteristic mesopause cooling was observed in airglow data (VDSR) Thus there is no obvious connection between the large-amplitude semidiurnal variation in the data we have presented and the sudden warming However, VDSR observed an enhancement in semidiurnal amplitudes after the mesospheric cooling Processes associated with sudden warmings can affect tides by redistributing ozone and modifying the thermal drive, or by changing the background wind system and modifying mode coupling[18,19] or the transmission of vertically propagating gravity waves

Neither thermal driving nor mode coupling seems viable because neither mechanism seems able to explain the amplification of the very-high-order modes required to explain large-amplitude tides at very high latitudes The required amplification of these modes cannot be due to ozone redistribution at polar latitudes because of the absence of sunlight, or to mode coupling because this process is ineffective in generating very-high-order modes from the lower-order energy-containing ones[19] Moreover, there is no reason to believe that mode coupling is significantly stronger in the pre-warming atmosphere than in the post-warming one In addition, it is very unlikely that changes in the thermal drive and mode coupling would act to amplify the semidiurnal tide and not the diurnal one No such amplification of the diurnal oscillation is observed at Spitsbergen (VDSR)

The observed enhancement of the semidiurnal oscillation might be due to an enhanced flux of gravity waves associated with the breakdown of the polar vortex An enhanced prevalence of large-amplitude waves has been observed (C R Philbrick, personal communication) and the semidiurnal pseudotide forced by this mechanism would be favoured over a one

We thank the NSF for financial support under research grants ATM85-42125 and ATM84-12982 This work was also supported in part by the Aerospace Sponsored Research Program

Received 23 June, accepted 1 October 1986

1 Walterscheid, R. L , DeVore, J G & Venkateswaran, S V *J atmos Sci.* **37,** 455-470 (1980)
2 Forbes, J M *J geophys Res* **87,** 5241-5252 (1982)
3 Walterscheid, R L *J geophys. Res* **86,** 9698-9706 (1981)
4 Mies, F H *J molec. Spectrosc.* **53,** 150-188 (1974)
5 Kvifte, G *Geofis Publs* **20,** 1-17 (1959)
6 Coxon, J A & Foster, S C *Can. J Phys* **60,** 41-48 (1982)
7 Krassovsky, V I *Ann Geophys* **28,** 739-746 (1972)
8 Myrabø, H K *Planet Space Sci.* **37,** 249-255 (1984)
9 Flattery, T W thesis, Univ Chicago (1967)
10 Longuet-Higgins, M S *Phil. Trans R. Soc.* **A262,** 511-607 (1968)
11 Vincent, R A. *J atmos terr Phys* **46,** 119-128 (1984)
12 Mason, A H , Gregory, J B & Stephenson, D G *J atmos Sci.* **32,** 1682-1688 (1975)
13 Hirota, I in *Dynamics of the Middle Atmosphere* (eds Holton, J R & Matsuno, T) 65-75 (Terra Scientific, Tokyo, 1984)
14 Vincent, R. A *J atmos terr Phys* **46,** 961-974 (1984)
15 Lysenko, I A , Orlyanskii, A. D & Portnyagin, Y I In *Coll. ext. Summ. Contrib Joint Assembly Seattle, Washington, CMUA Sess IAGA/IMAP, 22 August to 3 September 1977,* 31-1-31-7 (NCAR, Boulder, Colorado, 1977)
16 Manson, A H , Meek, C E & Labitzke, K *Adv Space Res* **4** (4), 41-55 (1984)
17 Quiroz, R. S *J geophys Res* **91** (in the press)
18 Lindzen, R. S & Hong, S -S *J atmos Sci.* **31,** 1421-1446 (1974)
19 Walterscheid, R L & Venkateswaran, S V *J atmos Sci.* **36,** 1623-1635 (1979)

Perturbation of the background radiation by a moving gravitational lens

L. I. Gurvits & I. G. Mitrofanov

Space Research Institute, Academy of Sciences USSR,
Profsojusnaja 84/32, 117810 Moscow, USSR

Birkinshaw and Gull[1] considered the perturbation of the cosmic background radiation by a moving gravitational lens (GL). An estimate of the maximum amplitude makes it seem possible that this effect may soon be observed, leading to the measurement of non-Hubble velocities of cosmological objects. Such measurements may become a significant test for cosmological models. We present here the general formula for this effect, which differs essentially from that in ref. 1.

Brightness $B(\nu, \mathbf{n})$ is transformed from a reference frame (A) to another (C), moving with velocity $\mathbf{V}_{AC}$ as follows [2]

$$B_C(\nu, \mathbf{n}_C) = (1-\Delta_{AC})^3 B_A[\nu/(1-\Delta_{AC}), \mathbf{n}_A] \qquad (1)$$

where

$$\Delta_{AC} = 1 - \frac{(1-V_{AC}^2/c^2)^{1/2}}{1-\mathbf{V}_{AC} \cdot \mathbf{n}_C/c} \simeq -\frac{\mathbf{V}_{AC} \cdot \mathbf{n}_C}{c}$$

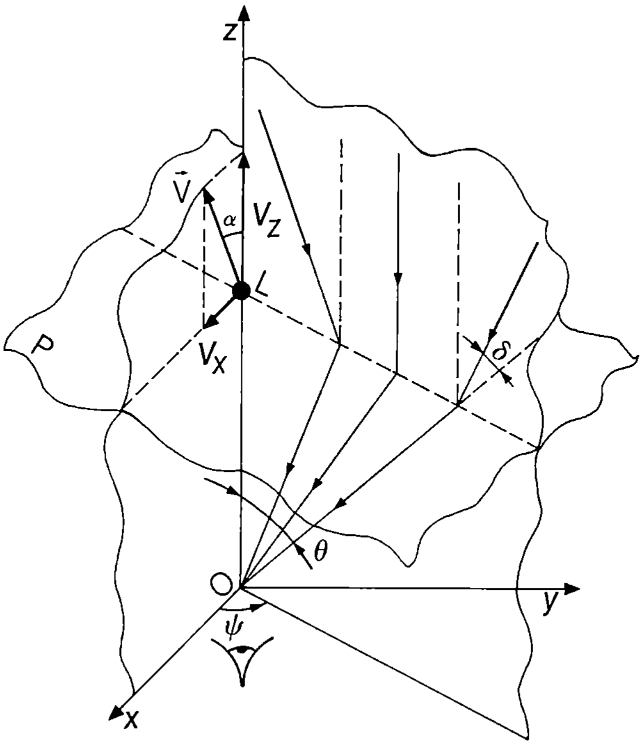


Fig 1 The geometry of the problem O is the observer, L is the position of the gravitational lens moving with velocity **V** in the X-Z plane, P is the picture plane, α is the angle between **V** and the Z axis, δ is the deflection angle, ψ is the position angle, and θ is the angle between the Z axis and the line of sight

ν is the radiation frequency in frame C, and $\mathbf{n}_A$ and $\mathbf{n}_C$ are the directions of propagation of the radiation at A and C, respectively To find the background brightness after lensing by a moving mass one uses transformation (1) twice the first time to transform from the background co-moving frame to the frame of a moving GL, and the second time to return from the moving GL frame to the background frame The deflection of geodetic lines in the GL frame is a cause of the perturbation in background brightness

Using this procedure one obtains for $\beta = V/c \ll 1$ the brightness perturbations

$$\frac{\Delta B(\nu, \mathbf{n})}{B_0(\nu)} = 2\beta[\cos\alpha \sin(\theta - \tfrac{1}{2}\delta)\sin\tfrac{1}{2}\delta - \sin\alpha\cos\psi\cos(\theta-\tfrac{1}{2}\delta)\sin\tfrac{1}{2}\delta] \times \frac{(h\nu/kT)\exp(h\nu/kT)}{\exp(h\nu/kT)-1} \qquad (2)$$

where B_0 and T are the initial background brightness, temperature and angles α, δ, ψ and θ are defined in Fig 1

In the Rayleigh-Jeans (RJ) part of a black-body spectrum the last factor of (2) is about 1 However, for the Wien part it is proportional to the frequency, but the brightness is exponentially small In the square brackets the first item corresponds to the longitudinal motion of GL It agrees with the previous result[3] in the RJ spectral region On the other hand, in ref 3 the 'transverse' part of the effect, corresponding to the second item of (2), was ignored For $\delta \ll \theta \ll 1$ the 'transverse' effect dominates the 'longitudinal' one It should be noted that transverse and longitudinal effects form different patterns in the picture plane The former does not depend on ψ, and thus the corresponding perturbation looks like a spot with circular symmetry, the latter includes the dipole term $\cos\psi$ and the corresponding perturbation has a central dipole asymmetry

Comparing (2) with equation (12) in ref. 1 one may conclude that the latter takes into account only the transverse effect. What is more, there is a discrepancy by a factor of -2 between amplitudes of the effect for the RJ spectral region. It seems that equation (12) in ref. 1 was obtained by excluding the factor $(1-\Delta_{AC})^3$ from the formula of brightness transformation such as (1).

An order of effect for the RJ spectrum may be estimated by a simple expression, which follows from (2),

$$\Delta B/B \sim \Delta T/T \sim \beta\delta \qquad (3)$$

Recently Collins *et al*[4] found indications of non-Hubble velocities of galaxies of the order of 1,000 km s^{-1}, i.e. $\beta \simeq 3\times10^{-3}$. One may use the angle between two lensed images to estimate the deflecting angle. There are some proposals (refs 5 and 6) that the two quasars 1146+111 B and C, separated by 157″, are in fact two lensed images. This value corresponds to $\delta \simeq 0.7\times10^{-3}$. Using these optimistic values we obtain $\Delta T/T \sim 2\times10^{-6}$. This is not far from today's sensitivity level, so the effect must be taken into account in future experiments.

We shall present elsewhere a general analysis of the complete effect based on this formula, both for $\delta < \theta$ and for $\delta \gg \theta$, and also for all β, not just for $\beta \ll 1$.

Received 13 May; accepted 22 August 1986.

1. Barkinshaw, M. & Gull, S. F. *Nature* **302**, 315-317 (1983).
2. Zel dovich, Ya. B. & Novikov, I. D. *The Structure and Evolution of the Universe* (University of Chicago Press, 1983).
3. Mitrofanov, I. G. *Pis'ma Astr. Zh.* (*Soviet Astr. Lett.*) **7**, 73-75 (1981).
4. Collins, C. A., Joseph, R. D. & Robertson, N. A. *Nature* **320**, 506-508 (1986).
5. Paczynski, B. *Nature* **319**, 567-568 (1986).
6. Turner, E. L. *et al. Nature* **321**, 142-144 (1986).

The near-parabolic flux and the origin of short-period comets

M. E. Bailey

Department of Astronomy, The University,
Manchester M13 9PL, UK

The origin of short-period comets, usually believed to arise by planetary perturbations acting on nearly parabolic orbits, has been unclear for many years[1]. Given the observed near-parabolic flux, the predicted number of such comets is too small by a factor of ~100 (refs 2–15). It has been suggested, therefore, that the adopted near-parabolic flux is in error[16], that a physical effect such as fragmentation[3,17,18] could operate, or that short-period comets do not come from the outer Oort cloud at all[9,10,13,19]. It has been proposed[13–15,20,21] that the problem could be resolved by postulating a comet belt or massive inner core to the Oort Cloud. Here I recalculate the flux of nearly parabolic orbits from the Oort Cloud, including the recently discovered dominant injection of comets by the galactic tide[22–25]. The observed number of short-period comets is consistent with the Oort Cloud having a massive inner core[26].

Consider a spherically symmetrical comet cloud characterized by an energy distribution $N(E)$ defined[26] so that the number of comets with energy per unit mass E in the range $(E, E+\mathrm{d}E)$ is

$$N(E)\,\mathrm{d}E = K(E/\hat{E})^{-\gamma}\,\mathrm{d}E \qquad (1)$$

where γ is the index of the energy spectrum ($\gamma = 5/2$ for Oort's standard model[27]), $E = -GM_\odot/2a$ is the specific energy of an orbit with semi-major axis a, $\hat{E} = -GM_\odot/2R_0$, $R_0 \simeq 2\times10^5$ AU is the outer radius of the cloud and K is a constant defined[26] in terms of the observed near-parabolic flux and other parameters of the model. One can assume that comets with perihelion distances $q < q_{\mathrm{lc}}$ are removed from the Oort Cloud by planetary perturbations, and adopt[28] $q_{\mathrm{lc}} \simeq 15$ AU. This defines a loss cylinder in velocity space, containing nearly parabolic comets with specific angular momenta $J < J_{\mathrm{lc}} = (2GM_\odot q_{\mathrm{lc}})^{1/2}$.

It is now possible to calculate the rate of injection of comets into the loss cylinder by stars and the galactic tide. The stellar injection rate per unit semi-major axis is[24,26]

$$F_{\mathrm{lc},*}(a) = 4\pi^2 f(a) J_{\mathrm{lc}}^2 \frac{GM_\odot}{2a^2}\begin{cases}\dfrac{\sigma_J^2(a)}{k^2(a)J_{\mathrm{lc}}^2} & a < a_* \\ 1 & a > a_*\end{cases} \qquad (2)$$

where a_* (see below) is the critical semi-major axis above which stellar perturbations alone are able to fill the loss cylinder, $f(a)$ is the phase-space distribution function defined in terms of equation (1) by[26]

$$f(a) = \frac{K}{8\pi^3} R_0^{-\gamma}(GM_\odot)^{-1/2} a^{\gamma-5/2} \qquad (3)$$

and $\sigma_J(a)$ is the r.m.s. change in specific angular momentum per revolution due to stellar perturbations. $k(a)$ is a slowly varying logarithmic function on the order of 2, which can be taken[24] to be

$$k^2(a) = \tfrac{1}{2}[\ln(a/2q_{\mathrm{lc}}) - 1] \qquad (4)$$

and[28]

$$\sigma_J^2(a) \simeq \tfrac{4}{3} P(a)\langle r^2 \varepsilon_*(r)\rangle \qquad (5)$$

where $P(a) = 2\pi(GM_\odot)^{-1/2}a^{3/2}$ is the orbital period and $\varepsilon_*(r)$ is the mean stellar energy transfer rate to comets at r with semi-major axes a (Heisler and Tremaine's expression[24] is not used because, as they remark, their result is invalid for $a \leqslant 4\times10^4$ AU.) The mean stellar energy transfer rate is given approximately by[29]

$$\varepsilon_*(a) \simeq \frac{4\pi G^2 M_*^2 n_*}{V_*}\begin{cases}(a/a_{\mathrm{c}})^2 & a < a_{\mathrm{c}} \\ 2\ln(a/a_{\mathrm{c}}) + 1 & a > a_{\mathrm{c}}\end{cases} \qquad (6)$$

where $a_{\mathrm{c}} \equiv \sqrt{\tfrac{12}{7}}\, b_{\min}$ and $b_{\min} = (2\pi n_* V_* T)^{-1/2}$ is the smallest impact parameter expected during the time T over which $\varepsilon_*(a)$ is averaged. Thus,

$$\frac{\sigma_J^2(a)}{k^2(a)J_{\mathrm{lc}}^2} = \frac{80\pi^2}{3}\,\frac{(GM_\odot)^{1/2}}{q_{\mathrm{lc}}}\,\frac{n_* M_*^2}{V_* M_\odot^2}\, a^{7/2} g(a) \qquad (7)$$

where

$$g(a) = \frac{1}{[\ln(a/2q_{\mathrm{lc}}) - 1]}\begin{cases}(a/a_{\mathrm{c}})^2 & a < a_{\mathrm{c}} \\ 2\ln(a/a_{\mathrm{c}}) + 1 & a > a_{\mathrm{c}}\end{cases} \qquad (8)$$

This allows the stellar-induced flux into the loss cylinder and a_* to be evaluated. Being concerned with filling the loss cylinder during one revolution, it is necessary to set $T = P(a)$, and adopt $n_* = 0.1$ pc^{-3}, $V_* = 30$ km s^{-1} and $\langle n_* M_*^2\rangle = \tfrac{4}{3}\times 0.03 M_\odot^2$ pc^{-3}, where the factor $\tfrac{4}{3}$ is to include the effects of close binaries[30]. This leads to $a_* \simeq 5.0\times10^4(q_{\mathrm{lc}}/15\ \mathrm{AU})^{2/7}$ AU, significantly larger than the value previously obtained[24]. This strengthens the conclusion[23,24] that stellar perturbations are relatively ineffective at filling the loss cylinder. (The results are different because with $T = P(a)$, $b_{\min}$ is typically in the range 1–2 × 10^4 AU and $g(a) \leqslant 1/2$. For comparison, if one sets $T = T_{\mathrm{SS}} = 4.5\times10^9$ yr, then $b_{\min} \simeq 700$ pc $\ll a$ and $g(a) \geqslant 1$. a_* is then $3.6\times10^4(q_{\mathrm{lc}}/15\ \mathrm{AU})^{2/7}$ AU, in agreement with the earlier value[24].) A similar result for a_* has been independently obtained by Fernández and Ip[31].

Next, consider the injection rate of comets by the galactic tide. This reduces[24] to

$$F_{\mathrm{lc,t}}(a) = 4\pi^2 f(a) J_{\mathrm{lc}}^2 \frac{GM_\odot}{2a^2}\begin{cases}\dfrac{10}{3\sqrt{2}}\,\dfrac{C_{\mathrm{gal}}}{GM_\odot q_{\mathrm{lc}}^{1/2}}\, a^{7/2} & a < a_{\mathrm{t}} \\ 1 & a > a_{\mathrm{t}}\end{cases} \qquad (9)$$

where a_t, the critical semi-major axis above which the tide just fills the loss cylinder, is given by

$$a_t = \left[\frac{3\sqrt{2}}{10} GM_\odot q_{lc}^{1/2} C_{gal}^{-1}\right]^{2/7} \simeq 3.3\times 10^4 \left(\frac{q_{lc}}{15\ \mathrm{AU}}\right)^{1/7}\ \mathrm{AU} \quad (10)$$

Here a galactic force law perpendicular to the galactic plane of the form $g_z = -C_{gal}z$, with $C_{gal} \simeq 10^{-29}\ \mathrm{s}^{-2}$ has been assumed[32]. This corresponds to a mid-plane density (see refs 24, 33) $\rho_0 = C_{gal}/4\pi G = 0.176 M_\odot \mathrm{pc}^{-3}$. For $a < a_t$ the ratio of the tidal to stellar injection rates is therefore

$$R_{t/*} = \frac{\sqrt{2}}{16\pi^2} \frac{q_{lc}^{1/2} V_*}{(GM_\odot)^{3/2}} \frac{C_{gal}}{\frac{1}{M_\odot^2}\langle n_* M_*^2\rangle} \frac{1}{g(a)} \quad (11)$$

which, with the adopted values of the parameters, reduces to $R_{t/*} = 1.933/g(a)$. Because $g(a) \simeq 1/5$ for $a = a_c \simeq 2\times 10^4$ AU one can conclude that the galactic tide dominates the stellar injection rate by about an order of magnitude at this value of a. For smaller a values, the tidal dominance of the injection rate increases.

This result (and that for a_* above) relies on equation (2), which in turn depends on a diffusion approximation to describe the quasi-steady distribution of comets near the loss cylinder. For $a \leq 2\times 10^4$ AU this assumption begins to break down because b_{min} becomes greater than a. Thus, although this approach should correctly give the general trend, the results are not expected to be precise. A recent Monte Carlo simulation by Heisler *et al*[34] indicates that the galactic tide dominates the stellar injection rate by a factor of ~4 at $a = 2\times 10^4$ AU.

An approximate total injection rate may thus be obtained by considering only the galactic tide. If a_0 is the semi-major axis corresponding to the inner edge of the Oort cloud, then

$$N_{tot} \simeq N_{tot,t} = \int_{a_0}^{R_0/2} F_{lc,t}(a)\, da \quad (12)$$

Substituting for f from (3) and introducing $x = a/a_t$, $\alpha_R = R_0/2a_t \gg 1$ and $\alpha_0 = a_0/a_t \ll 1$, this reduces to

$$N_{tot} = \frac{KR_0^{-\gamma}}{2\pi}(GM_\odot)^{3/2} q_{lc} a_t^{\gamma-7/2}\left\{\int_1^{\alpha_R} x^{\gamma-9/2}\, dx + \int_{\alpha_0}^1 x^{\gamma-1}\, dx\right\} \quad (13)$$

In this expression, the first term corresponds to comets with almost parabolic orbits and a nearly uniform distribution of perihelion distances, whereas the second corresponds to comets with rather shorter periods and perihelia close to q_{lc} (for example ref 28). The comets in this second group cannot be directly observed (as $q_{lc} \simeq 15$ AU), but those in the first group can be detected up to distances of the order of 5 AU. This allows the leading term of (13) to be re-written in terms of the observed nearly parabolic flux F_{obs}, defined here as the number of comets passing perihelion per unit time with perihelia less than some value D_{obs}. The ratio F_{obs}/D_{obs} is then nearly constant. Thus, from (13),

$$F_{obs} = \frac{KR_0^{-\gamma}}{2\pi}(GM_\odot)^{3/2} q_{lc} a_t^{\gamma-7/2} \int_1^{\alpha_R} x^{\gamma-9/2}\, dx \frac{D_{obs}}{q_{lc}} \quad (14)$$

so

$$N_{tot} = \frac{F_{obs}}{D_{obs}} q_{lc}\left\{1 + \frac{\int_{\alpha_0}^1 x^{\gamma-1}\, dx}{\int_1^{\alpha_R} x^{\gamma-9/2}\, dx}\right\} \quad (15)$$

The total number of short-period comets may now be obtained[13-15] by multiplying N_{tot} by three factors: the fraction f_{ι_0} of comets in the loss cylinder with small enough inclinations ($0 \leq \iota \leq \iota_0$) to be significantly affected by planetary perturbations, the mean planetary capture probability for each group of incoming comets including the effects of all the planets, and the average lifetime T_{SP} of observable short-period comets. The first factor may be approximately estimated by assuming that the incoming comets have random inclinations. In this case, although the assumption is not strictly accurate when the galactic tide drives the influx, $f_{\iota_0} = \frac{1}{2}(1 - \cos \iota_0)$. The second factor may be written as an integral over perihelion distance of the mean planetary capture probability $p(q)$, weighted according to the perihelion distribution of each group of comets. For the first group (distributed uniformly with q) this becomes

$$\bar{p} = \frac{1}{q_{lc}} \int_0^{q_{lc}} p(q)\, dq \quad (16)$$

whereas for the second group (concentrated close to $q = q_{lc}$) it is simply $\bar{p} = p(q_{lc})$. Following Fernández[13] in assuming that the mean planetary capture probability $p(q)$ has the step-wise form

$$p(q) = \begin{cases} 1/128 & 0.0 < q \leq 5.8\ \mathrm{AU} \\ 1/700 & 5.8 < q \leq 10.7\ \mathrm{AU} \\ 1/3{,}200 & 10.7 < q \leq 22\ \mathrm{AU} \\ 1/6{,}000 & 22 < q \leq 34\ \mathrm{AU} \end{cases} \quad (17)$$

gives $\bar{p} \simeq 0.054\ \mathrm{AU}/q_{lc}$ for the first group, and $\bar{p} = 1/3{,}200$ or $1/700$ for the second group, depending respectively on whether the planetary captures are dominated by Uranus ($q_{lc} = 15$ AU) or Saturn ($q_{lc} = 10$ AU). Adopting the values for Uranus, and setting[14,15] $\iota_0 = 30°$, $T_{SP} = 3{,}000$ yr, and $F_{obs} = 0.7\ \mathrm{yr}^{-1}$, corresponding to $D_{obs} = 2.5$ AU, then

$$N_{SP} \simeq \frac{F_{obs}}{D_{obs}} f_{\iota_0} T_{SP}\left\{\int_0^{q_{lc}} p(q)\, dq + p(q_{lc}) q_{lc} \frac{\int_{\alpha_0}^1 x^{\gamma-1}\, dx}{\int_1^{\alpha_R} x^{\gamma-9/2}\, dx}\right\} \quad (18)$$

or

$$N_{SP} \simeq 56.3\left\{0.054 + \frac{15}{3200}F(\gamma)\right\} \quad (19)$$

where $F(\gamma)$ is the ratio of the integrals appearing in the second term of (18). The observed number of short-period comets (~10^2) can thus be understood if $F(\gamma) \simeq 200$. This is clearly not possible in Oort's standard model with $\gamma = 5/2$, but for arguably more plausible models[26,30,35] with $\gamma < 0$ it is certainly possible. For example, adopting $R_0 = 2\times 10^5$ AU, $a_t = 3\times 10^4$ AU, $a_0 = 4\times 10^3$ AU and $\gamma = -2$, then we find $F(\gamma) \simeq 152$ and $N_{SP} \simeq 43$. On this model the flux of long-period comets through the Uranus capture zone is more than two orders of magnitude greater than expected from a simple linear extrapolation of the observed local flux.

A check on the model is provided by the total number of comets. This is readily calculated[26] as

$$N_0 = KR_0^{-\gamma}\frac{GM_\odot}{2}\int_{a_0}^{R_0/2} a^{\gamma-2}\, da$$

that is,

$$N_0 = \frac{F_{obs}}{D_{obs}}\pi(GM_\odot)^{-1/2} a_t^{5/2} \frac{\int_{\alpha_0}^{\alpha_R} x^{\gamma-2}\, dx}{\int_1^{\alpha_R} x^{\gamma-9/2}\, dx} \quad (20)$$

For the above parameters this gives $N_0 = 1.7\times 10^{13}$. With a mean mass per comet[36] ~2.6×10^{14} kg, the total mass of the cloud becomes $\simeq 700 M_\oplus$, a large but not unreasonable amount[37]. It is possible that such a mass of comets might already have been detected by the Infra-Red Astronomy Satellite[38-40].

Thus I have shown that the observed number of short-period comets may be understood in terms of the capture of Oort Cloud comets by planetary perturbations, provided that the cloud has a massive inner core and that the injection of new comets is dominated by the galactic tide. This is because the ratio of the tidal injection rate to that of the stars increases sharply ($\propto a^2$) for semimajor axes $\lesssim 3\times 10^4$ AU, and although injection by the galactic tide is less effective at smaller semi-major axes (equation (9)), the relative decline may be more than off-set if there is a sufficiently large number of comets in the inner core. As with some previous suggestions for overcoming the short-period comet problem[9,10,14,19], this analysis predicts that most short-period comets are initially captured by Uranus (or possibly Saturn), rather than Jupiter In particular, the present theory predicts that most short-period comet progenitors are injected by the galactic tide with relatively small semi-major axes close to a_0 ($\simeq$5,000 AU) In this case the short-period comet population would bear little or no relation to the observed near-parabolic flux, and the usual links between these comets and the more numerous observed long-period group are brought into question These results indicate the importance of detailed Monte Carlo simulations[41] aimed at discovering whether such original orbits evolve into observable short-period comets

I thank S V M Clube and W M Napier for comments on the manuscript and S D Tremaine for helpful suggestions and a preprint[34] This work was supported by the SERC

Received 31 July, accepted 1 October 1986

1 Bailey, M E, Clube, S V M & Napier, W M *Vistas in Astr* **29**, 53-112 (1986)
2 Proctor, R A *Knowledge* **6**, 126-128 (1884)
3 Bredichin, T *Ann. de l'Obs de Moscou* (Ser 2) **2**, 1-17 (1889)
4 Russell, H N *Astr J* **33**, 49-61 (1920)
5 Proctor, M & Crommelin, A C D *Comets* (Technical Press, London, 1937)
6 Everhart, E *Astr J* **74**, 735-750 (1969)
7 Havnes, O *Icarus* **12**, 331-337 (1970)
8 Everhart, E *Astrophys Lett* **10**, 131-135 (1972)
9 Vaghi, S *Astr Astrophys* **24**, 107-110 (1973)
10 Joss, P C *Astr Astrophys* **25**, 271-273 (1973)
11 Kazimirchak-Polonskaya, E I *IAU Symp* **45**, 373-397 (1972)
12 Everhart, E *IAU Colloq* **39**, 99-104 (1977)
13 Fernandez, J A *Mon Not R astr Soc* **192**, 481-491 (1980)
14 Fernandez, J A & Ip, W-H *Asteroids, Comets, Meteors* (eds Lagerkvist, C I & Rickman, H) 387-390 (Uppsala Observatory, Uppsala, 1983)
15 Fernandez, J A & Ip, W-H *Icarus* **54**, 377-387 (1983)
16 Hughes, D W *IAU Colloq* **83**, 129-142 (1985)
17 Clube, S V M & Napier, W M *Quart Jl R astr Soc* **23**, 45-66 (1982)
18 Napier, W M *Asteroids, Comets, Meteors* (eds Lagerkvist, C I & Rickman, H) 391-395 (Uppsala Observatory, Uppsala, 1983)
19 Delsemme, A H *Astr Astrophys* **29**, 377-381 (1973)
20 Mendis, D A *Astrophys Sp Sci* **20**, 165-176 (1973)
21 Bailey, M E *Asteroids, Comets, Meteors* (eds Lagerkvist, C-I & Rickman, H) 383-386 (Uppsala Observatory, Uppsala, 1983)
22 Torbett, M V *The Galaxy and the Solar System* (eds Matthews, M, Smoluchowski, R & Bahcall, J N) (University of Arizona Press, Tucson, in the press)
23 Torbett, M V *Mon Not R astr Soc* (in the press)
24 Heisler, J & Tremaine, S D *Icarus* **65**, 13-26 (1986)
25 Morris, D E & Muller, R A *Icarus* **65**, 1-12 (1986)
26 Bailey, M E *Mon Not. R astr Soc* **204**, 603-633 (1983)
27 Oort, J H *Bull Astr Inst Neth* **11**, 91-110 (1950)
28 Bailey, M E *Mon. Not R astr Soc* **211**, 347-368 (1984)
29 Bailey, M E *Asteroids, Comets, Meteors II* (eds Lagerkvist, C-I, Lindblad, B A, Lundstedt, H & Rickman, H) 207-210 (Uppsala University, Uppsala, 1986)
30 Bailey, M E *Mon. Not R astr Soc* **218**, 1-30 (1986)
31 Fernandez, J A & Ip, W-H Preprint
32 Wielen, R *Astr Astrophys* **25**, 285-297 (1973)
33 Bahcall, J N & Bahcall, S *Nature* **316**, 706-709 (1985)
34 Heisler, J, Tremaine, S D & Alcock, C Preprint
35 Hills, J G *Astr J* **86**, 1730-1740 (1981)
36 Fernández, J A *Astr J* **87**, 1318-1332 (1982)
37 Bailey, M E *Nature* **302**, 399-400 (1983)
38 Rowan-Robinson, M *Mon Not R astr Soc* **219**, 737-749 (1986)
39 Bailey, M E *Mon. Not R astr Soc* **205**, 42P-25P (1983)
40 Bailey, M E, McBreen, B & Ray, T P *Mon Not R astr Soc* **209**, 881-890 (1984), *Corrigendum* **211**, 255
41 Rickman, H & Vaghi, S *Astr Astrophys* **51**, 327-342 (1976)

Inferences on the evidence for radioactive ^{53}Mn in the early Solar System

Typhoon Lee

Institute of Earth Sciences, Academia Sinica, PO Box 23-59,
Taipei, Taiwan (R O C) 10769

Several now-extinct radionuclides are believed to have been present in the early Solar System[1] and have been used to infer the time-scales for various processes during its formation. However, the precise interpretation of the chronometric significance of an extinct nuclide depends on whether isotopic heterogeneities exist in that nuclide and/or in its daughter product. This is probable because such heterogeneities have already been found in many elements. Birck and Allègre[2] have recently reported data suggesting the presence of a new extinct nuclide, ^{53}Mn (mean life $\tau = 5$ Myr). In this letter I show that their data are inconsistent with the prediction of a single-stage model in which both Mn and its daughter element Cr were isotopically homogeneous. I also explore the virtues and consequences of the three alternative interpretations of the inconsistency: (1) the host material has experienced multiple-stage evolution, (2) ^{53}Mn was heterogeneously distributed, or (3) ^{53}Cr has intrinsic anomalies.

The discovery of ^{26}Al, whose mean life is shorter than that of ^{53}Mn, in the early Solar System suggests the possibility that ^{53}Mn may have also existed[3] Recent searches for ^{53}Mn began with that of Nitoh *et al*[4] Lee and Tera[5 6] placed more stringent upper limits on its abundance However, positive results were not obtained until Birck and Allègre[2] developed superior analytical techniques enabling high-precision isotopic measurement of Cr (ref 7) They also discovered isotopic anomalies in ^{54}Cr in ordinary Allende inclusions, which were confirmed by Papanastassiou *et al*[8] Papanastassiou[9] further discovered much larger ^{54}Cr and ^{53}Cr anomalies in FUN (fractionation and unknown nuclear) inclusions The discovery of anomalies intrinsic to the Cr isotopes accentuates the importance of considering the effects of Mn and Cr isotopic heterogeneities in the discussion of the reported evidence for the presence of ^{53}Mn

Figure 1*a* shows the systematics of ^{53}Mn–^{53}Cr evolution in a diagram of ^{53}Cr/^{52}Cr against ^{55}Mn/^{52}Cr The time (before present) evolution of ^{53}Cr/^{52}Cr due to ^{53}Mn decay for object P formed at t_1 from source S can be described by

$$({}^{53}\mathrm{Cr}/{}^{52}\mathrm{Cr})_P(t) = I_1 + ({}^{53}\mathrm{Mn}/{}^{55}\mathrm{Mn})\,(t_1)\,(1 - \exp(t - t_1)/\tau) \times ({}^{55}\mathrm{Mn}/{}^{52}\mathrm{Cr})_P \qquad (1)$$

provided that this object and its source share the same initial ^{53}Cr/^{52}Cr(t_1) ($\equiv I_1$) and ^{53}Mn/^{55}Mn(t_1) and that the object is a closed system against Mn and Cr transport Notice that equation (1) represents a straight line in Fig 1*a* connecting the object, the source and the initial, whose slope increases with time up to the initial Mn ratio, whereas its intercept remains fixed at the initial Cr ratio However, at any time, the line always passes through the source composition for that time (heavy lines in Fig 1*a*) A group of such objects formed simultaneously at t_1 with initially homogeneous Mn and Cr ratios will thus form a linear array This is the well-known isochron for an extinct nuclide If such a group, after forming from the source, has remained closed until now (that is, experienced only single-stage evolution) then an isochron would be observed today and taken as evidence for the presence of ^{53}Mn at the formation time of the group I have thus shown that, in a single stage evolution model with initial isotopic homogeneities, the isochron always passes through the source point Applying this to the Allende meteorite and considering either different inclusions or different minerals in a single inclusion as a group formed directly from the solar nebula, we would thus expect that all such isochrons should pass through the normal Solar System point

Birck and Allègre[2] reported two 'isochrons', one of which consists of whole-rock data for the Allende inclusion with a slope corresponding to an initial ^{53}Mn/^{55}Mn ratio of $6\ 7 \times 10^{-5}$ The other consists of whole rock data for Allende inclusions with a slope corresponding to an initial ^{53}Mn/^{55}Mn ratio of $3\ 7 \times 10^{-5}$ Figure 1b shows these isochrons and also some key data points The whole inclusion isochron hinges on the data for G6 as noted in ref 2, also, it does not pass through SB The mineral isochron data span a relatively narrow range in Mn/Cr, giving a difference in ^{53}Cr/^{52}Cr only six times the standard error Further Cr isotopic studies, especially on primitive high-Mn/Cr samples, appear necessary to substantiate these ^{53}Mn isochrons

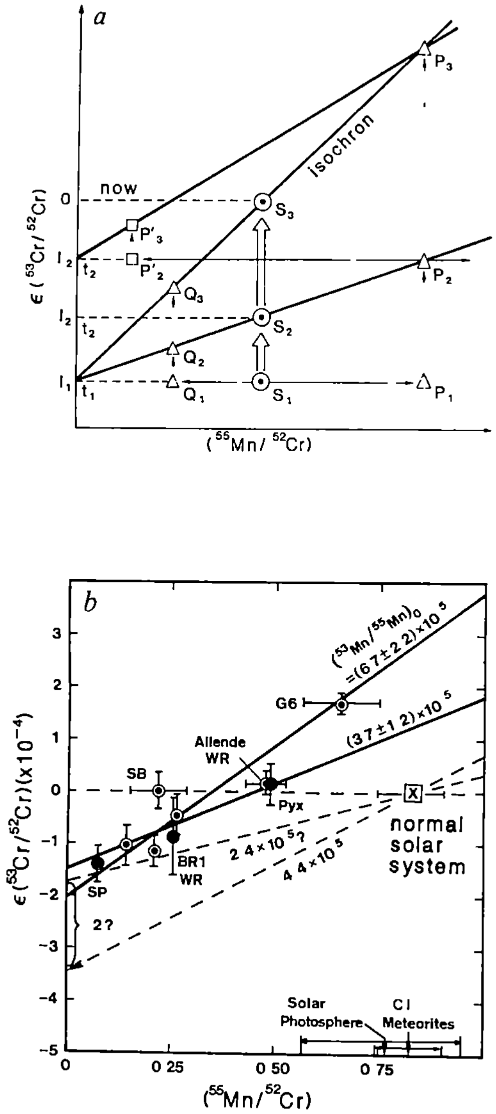


Fig 1 Mn-Cr isochron diagram $\varepsilon(^{53}Cr/^{52}Cr)$ is the fractional deviation of $^{53}Cr/^{52}Cr$ from the normal in unit of 0 01% *a*, Schematics The evolution of the Cr ratios by ^{53}Mn decay for a group of objects (*P*, *Q*, etc) formed from source *S* at time t_1 is described by equation (1) if they share the same initial Cr and Mn isotopic ratios as the source at t_1 and if they remain a close system They plot as a linear array ('isochron') whose intercept is fixed at the initial Cr ratios I_1 and whose slope gradually steepens, eventually reaching the initial $^{53}Mn/^{55}Mn$ ratio At any time the isochron for a group directly formed from the source (single-stage evolution) always passes through the source composition at that time, so the tie lines I_iS_i always coincide with isochrons P_iQ_i for $i = 1, 2, 3$ at time t_1, t_2 and 'now' Therefore, in this single-stage homogeneous model, the isochrons for Allende inclusions are expected to pass through the putative source composition, that of the Solar System (⊙) In contrast, the Solar-System point may show an offset from the isochrons in a multiple-stage model For example, the isochron for a group of objects like *P'*, which formed from an intermediate source reservoir P_2 at time t_2 that ultimately formed from *S* at an earlier time, does not pass through *S* The direction of the offset depends on whether the Mn/Cr for *P* is higher or lower than that of *S* *b*, Observed isochrons[2] and interpretations The observed isochrons do not pass through the Solar-System point and the uncertainty in the Solar-System Mn/Cr estimates based on either meteoritic or photospheric observations does not seem to be large enough to eliminate the offset If a two-stage evolution model for the Allende samples is invoked as the explanation, then the intermediate source reservoir must have a Mn/Cr higher than the solar value because of the direction of the offset Alternatively, Mn isotopic heterogeneity can explain the offset if the solar reservoir has a $^{53}Mn/^{55}Mn = 2\ 4 \times 10^{-5}$ assuming that it and the samples started with the same initial $^{53}Cr/^{52}Cr$ The offset can also be explained if the initial $^{53}Cr/^{52}Cr$ for the Allende inclusions is higher than that of the normal solar reservoir by about 0 02% The ratio of the inferred intrinsic ^{53}Cr excess to the observed ^{54}Cr excess (~0 3) for these ordinary inclusions is identical to that measured in the unusual FUN inclusion EK1-4-1 (ref 9) and thus may have come from a single anomalous component of Cr

In Fig 1*b* I have also plotted the normal Solar System point at $\varepsilon(^{53}Cr/^{52}Cr) = 0$ and $(^{55}Mn/^{52}Cr) = 0\ 85$ (ref 10) The 'isochrons' do not pass through the Solar System points, which contrasts with the prediction of the above model

We must first consider the possibility that the currently accepted Mn/Cr ratio for the Solar System is in error The spectroscopically determined solar photospheric value is 0 60 (corresponding to a $^{55}Mn/^{52}Cr$ ratio of 0 76) with an uncertainty of ~25% (ref 11) This estimate is consistent with the 'cosmic' value of 0 71, based on type CI meteorites, which has a nominal uncertainty of ~13% (see ref 10 and references therein) However, this is based almost exclusively on the data of Orgueil, whereas some of the limited data on Alais are discrepant An Mn/Cr of 0 3 is required to shift the Solar System point to the isochron If this is the correct interpretation, our present estimate for the Solar System Mn/Cr has to be too high by about a factor of two This seems unlikely in view of the consistency between the two estimates using different methods and their associated errors

If the isotopic measurements and the estimated solar Mn/Cr value are correct then one or more of the three alternatives mentioned earlier must explain the offset of the Solar System point Although the combination of more than one cause is possible it would be too intractable to allow a useful discussion, so I shall exclude the simple model's assumptions only one at a time The first alternative, multiple-stage evolution of the host material, is difficult to quantify This is because of the lack of constraints on the evolution during intermediate stages There are many combinations of the duration and the Mn/Cr ratio for the intermediate stage that can explain the reported offset Figure 1*a* shows a relatively simple two-stage evolution model that can be used to explain the offset and its direction Object P_1 formed at t_1 from S_1 would reach point P_2 at t_2, hence a sample formed from P_2, acting now as the intermediate source reservoir, would have the initial ratio I'_2, much higher than I_2, which is the initial for an object formed directly from S_2 Connecting the initial with the intermediate source, line $I'_2\ P_3$ is the isochron observed today for the samples that experienced the two-stage evolution It passes through P'_3 but not through S_3, thus producing an offset from the Solar System point We note that the direction of the offset depends on whether the intermediate-stage reservoir has a Mn/Cr ratio higher or lower than the solar value Since the reported isochrons are to the left of the Solar System point, the intermediate-stage reservoir must have a Mn/Cr ratio higher than the solar value This is somewhat surprising because in refractory samples such as Allende inclusions this ratio is lower than the solar value The required high-Mn/Cr intermediate source reservoir from which the inclusions formed must be less refractory than the inclusions

The remaining two alternatives have to do with isotopic heterogeneities rather than chemical fractionation The offset of the normal Solar System point from the isochrons can be understood if Allende inclusions have a higher $^{53}Mn/^{55}Mn$ than the Solar System However, because the initial $^{53}Cr/^{52}Cr$ of the inclusions is lower than normal and we assume that the initial Cr ratios of the inclusions were the same as the source (that is the solar) reservoir, the ^{53}Mn abundance in the normal Solar System material must be non-zero so that its decay can eventually increase the $^{53}Cr/^{52}Cr$ to the present value The required $^{53}Mn/^{55}Mn$ ratio can be obtained by connecting the Solar System point to the initial Cr ratio of the observed isochrons The slope of this line corresponds to a $^{53}Mn/^{55}Mn$ of about $2\ 4 \times 10^{-5}$ (see Fig 1*b*) We note that inclusion SB has a normal $^{53}Cr/^{52}Cr$ today If it also started from the same initial Cr ratio as other inclusions then it coincidentally started from the right amount of ^{53}Mn, higher than that for other inclusions, to give a normal $^{53}Cr/^{52}Cr$ today Or, more likely, it equilibrated later with the Solar System reservoir after all the ^{53}Mn had decayed away This situation is reminiscent of the case of ^{26}Al where heterogeneities were also suspected[1]

Table 1 Estimated composition of the anomalous component

Nuclide	Observed abundance*	e − process abundance† $\eta = 0\ 0710$	$\eta = 0\ 0837$	Normal
^{50}Cr	0	0 5	0 00001	1 8
^{52}Cr	0	419	2 7	36
^{53}Cr	1 3	1 2	0 18	4
^{54}Cr	1	1	1	1
^{53}Mn	0 7	16	0 003	—

* For assumptions see text
† Ref 12, all ^{54}Mn are assumed to have decayed into ^{54}Cr

The last alternative is that the Allende samples began with a ^{53}Cr/^{52}Cr higher than that for the normal solar system material The required amount can be found from Fig 1*b* If we draw an isochron through the Solar System point with a slope equal to the average of the published isochrons, this line would intercept the vertical axis at ~0 02% below the initial for the Allende samples Thus the offset can be explained if the inclusions started with an initial intrinsic ^{53}Cr excess of 0 02% relative to the Solar System ratio at the time of formation of the inclusions Because these samples already have excesses in ^{54}Cr of about 0 06%[7,8], it seems plausible that the neighbouring neutron-rich Cr isotope also has some accompanying excesses The possibility that part of ^{53}Cr effects in ordinary inclusions is intrinsic to this nuclide is enhanced by the recent discovery of a large ^{53}Cr excess in FUN inclusion EK1-4-1 by Papanastassiou[9], who concluded that it is more reasonble to interpret the large effect as the result of intrinsic variations in ^{53}Cr than by invoking the *in situ* decay of ^{53}Mn If so, and if both of the observed ^{53}Cr and ^{54}Cr effects in EK1-4-1 are caused by mixing an anomalous component with the normal, then the ratio of ^{53}Cr to ^{54}Cr in that component would be 0 16%/0 48% = 0 3 times the normal ratio After allowance is made for the contribution of ^{53}Mn effects, the ratio of the inferred intrinsic ^{53}Cr excess to the observed ^{54}Cr excess in ordinary inclusions is also (0 02%/0 06%) = 0 3 Therefore, if the intrinsic Cr anomalies inferred for the ordinary inclusions are also to be explained by the admixing of an anomalous component with the normal Solar System matter then a single component would be sufficient for both EK1-4-1 and ordinary inclusions

It is important to consider whether the ^{53}Mn, ^{54}Cr and the inferred ^{53}Cr were co-produced and represent a single 'anomalous' component We can estimate the composition of this component by assuming that it lacks ^{50}Cr and ^{52}Cr (ref 2) Using the data on inclusion BR1 (^{54}Cr excess = 0 048% (ref 2), ^{53}Mn/^{55}Mn = 37 p p m (ref 2) Mn/Cr = 0 21 (ref 2), and ^{53}Cr excess = 0 016%, inferred above) we arrive at the composition relative to ^{54}Cr given in the first column in Table 1 The desired composition seems to consist of comparable numbers of ^{54}Cr, ^{53}Cr and ^{53}Mn atoms but no ^{50}Cr and ^{52}Cr It would be interesting to explore if such a pattern can be produced in a single nucleosynthetic source As a first step, Table 1 shows a comparison of this composition with the production patterns given by Hainebach *et al*[12] as well as the normal It may indeed be possible to produce the desired composition in a single e-process zone characterized by η, the neutron excess parameter, somewhere between 0 071 and 0 0837, however, overproduction of ^{52}Cr could be a problem The production pattern is too sensitive to η to allow a meaningful interpolation between the two η values where the nucleosynthetic calculations were performed Because ^{53}Mn is neutron-poor and ^{53}Cr and ^{54}Cr are neutron-rich, one might attempt to assign these three nuclides to at least two different zones However, the observed ^{53}Mn abundance is small, so a single-zone model remains a possibility This particular e-process model is shown here merely to illustrate the usefulness of such an estimated composition and the possibility of producing it in a single zone Such a zone would produce Ti highly enriched in ^{50}Ti but more ^{46}Ca than ^{48}Ca Thus it cannot fully account for the rough qualitative correlation observed in Allende inclusions between ^{54}Cr, ^{50}Ti and ^{48}Ca Also, stellar nucleosynthesis considerations suggest that e-process products ejected from a massive star are expected to be a mixture of materials from zones covering a broad spectrum of η[13]

We have demonstrated that the offset of the Solar System point from the ^{53}Mn isochron reported in ref 2 disagrees with the prediction of a model in which Allende inclusions formed directly from the solar nebular with homogeneous initial ^{53}Mn/^{55}Mn and ^{53}Cr/^{52}Cr ratios This disagreement could mean that the inclusions have a more complicated history If they have formed from an intermediate source reservoir which in turn formed from the solar reservoir, then this intermediate source must have a higher Mn/Cr than the solar ratio Alternatively, the offset could be explained by heterogeneities in the distribution of ^{53}Mn and the inclusions had an initial ^{53}Mn twice the average Solar System value The third possibility is that when they were formed the inclusions had an intrinsic ^{53}Cr excess of 0 02% relative to normal Solar System matter A single component of anomalous Cr with a ^{54}Cr/^{53}Cr three times the normal ratio can explain not only this inferred ^{53}Cr excess and observed ^{54}Cr effects in ordinary inclusions but also the much larger Cr anomalies found in the unusual FUN inclusion EK1-4-1 If ^{53}Mn is also co-produced with this component then the nucleosynthesis of these iron-group nuclides would be much better constrained, and a single e-process zone would be a possible source

I thank Drs D A Papanastassiou (Caltech) and J L Birck (UCSD and Univ Paris) for comments that enabled me to improve the paper, Dr K K Liu of the Academia Sinica for useful discussions, and the National Science Council (Taiwan, R O C) for travel grants to meetings where the ideas in this paper were first formed

Received 19 March, accepted 19 August 1986

1 Wasserburg, G J *Proto-stars and Planets II* (ed Black, D C & Matthews, M S) 703-707 (Univ Arizona Press, 1985)
2 Birck, J L & Allegre, C J *Geophys Res Lett* **12**, 745-748 (1985)
3 Lee, T , Papanastassiou, D A & Wasserburg, G J *Geophys Res Lett* **3**, 109-112 (1976)
4 Nitoh, O , Kamata, M & Honda, M *Abstr Vol 4th Int Conf Geochron*, Snowmass, Colorado, 308-310 (1978)
5 Lee, T & Tera, F *Abstr Vol 5th Int Conf Geochron* Nikko, Japan, 212-213 (1982)
6 Lee, T & Tera, F *Geochim cosmochim Acta* **50**, 199-206 (1986)
7 Birck, J L & Allegre, C J *Geophys Res Lett* **11**, 943-946 (1984)
8 Papanastassiou, D A , Zyskind, J L , Ngo, H H & Wasserburg, G J *Meteoritics* **20** (in the press)
9 Papanastassiou, D A *Astrophys J Lett* (in the press)
10 Anders, E & Ebihara, M *Geochim cosmochim Acta* **46**, 2363-2380 (1982)
11 Grevesse, N in *Frontiers of Astronomy and Astrophysics* (ed Pallavicini, R) 71-82 (Ital Ast Soc , Florence, 1984)
12 Hainebach, K L , Clayton, D D , Arnett, W D & Woosley, S E *Astrophys J* **193**, 157-168 (1974)
13 Hartmann, D Woosley, S E & El Eid, M F *Astrophys J* **297**, 837 (1985)

EISCAT observations of the E-region semidiurnal tide

T. S. Virdi, G. O. L. Jones & P. J. S. Williams

University College of Wales, Department of Physics, Penglais,
Aberystwyth SY23 3BZ, UK

Many observers have reported a semidiurnal tide in the E-region of the ionosphere[1,7] Theory demonstrates the possibility of several tidal modes with the relative strength of each mode depending on the efficiency of generation in the stratosphere, the attenuation suffered in propagating through the mesosphere and mode-coupling in regions of nonlinearity. To determine which modes actually occur it is important to observe the tide at all heights in the E-region and at a range of latitudes. EISCAT, the European incoherent scatter radar[2,3], is an ideal facility to make such observations at 69.5° N At this latitude, measurements are usually

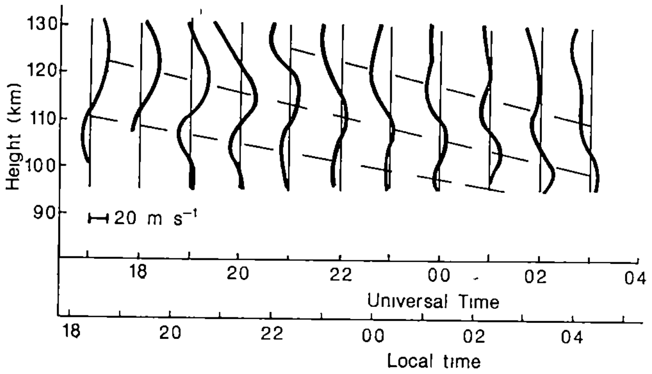


Fig. 1 Height profiles of ion velocity parallel to the magnetic field line, 14-15 February 1985

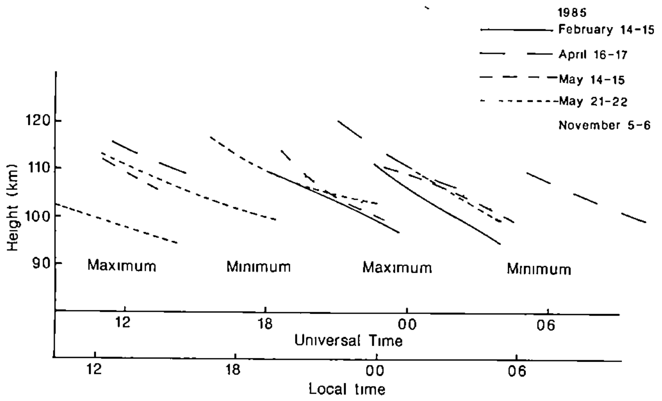


Fig 2 The descending phase-fronts of the semidiurnal tide as demonstrated by the maxima (upward) and minima (downward) of ion velocity parallel to the magnetic field above Tromsø on five days during 1985

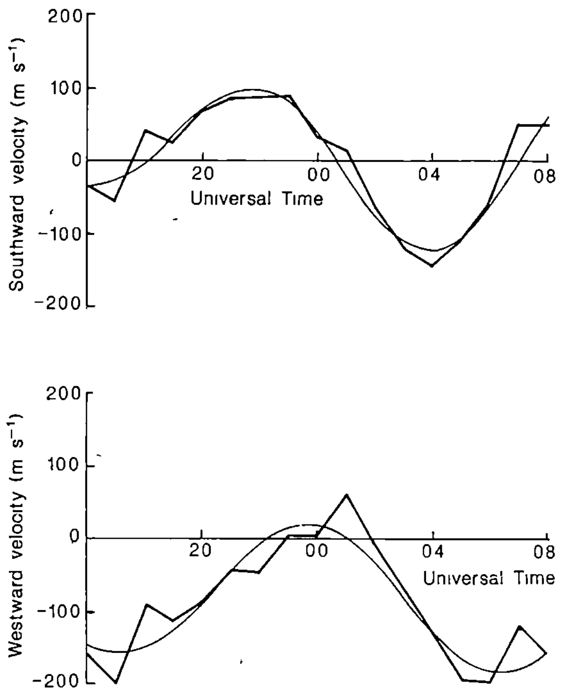


Fig 3 Horizontal velocity in the southward and westward direction at a height of 101 km, 5-6 November 1984 Heavy line, observed velocity, tight line, best-fitting diurnal and semidiurnal tide

perturbed by strong auroral heating, but near sunspot minimum auroral disturbances are sometimes very small and the underlying pattern of the semidiurnal tides can be observed. Such conditions occurred in 1985 and here we report observations on five different days that illustrated the dominance of the (2,4) mode at high latitudes.

At altitudes below 120 km in the ionosphere, the ion-neutral collision frequency is much greater than the ion-gyro frequency so that movement of the ion population is mainly controlled by the neutral atmosphere Measurements of ion velocity can therefore monitor neutral winds

The data presented here were taken while EISCAT was performing the Common Programme CP1 The antenna at Tromsø transmitted along the magnetic field line and observed the whole E-region with a height resolution of 3 km Remote antennas at Kiruna and Sodankyla observed the signals scattered from heights of 101, 110, 119, 131 and 312 km The data were analysed to provide height profiles of electron density, electron and ion temperature and plasma velocity

Figure 1 shows hourly averages of plasma velocity along the field line, $v_{\|}$, as a function of height during 14-15 Feburary A sinusoidal variation with height is observed, and this pattern descends steadily at a rate of 2-3 km h^{-1}

In Fig 2 the heights of the maxima and minima in $v_{\|}$ are plotted against time for each of the five days The pattern is consistent, and small differences from day to day can be accounted for by seasonal changes in the neutral atmosphere, day-to-day variations in the relative strength of the different diurnal and semidiurnal modes, the effects of auroral electric fields, especially above 120 km, and random errors in the measurements themselves

To determine the actual wind velocity associated with the tidal modes we calculated the full vector of plasma velocity, $\mathbf{v}_p$, by combining the separate components of velocity measured at the three EISCAT stations and applying a small correction for the electric field, derived from measurements at 312 km Figure 3 shows typical results the horizontal north-south and east-west velocities on 5-6 November are plotted against time for a height of 101 km, showing the semidiurnal tide

Theory[4,5] predicts that the tidal mode (m,n) shows a downward phase propagation with a vertical wavelength given by

$$\lambda_{z,mn} = 2\pi H / \sqrt{\frac{H}{h_{mn}}\left(\frac{\gamma-1}{\gamma}+\frac{\mathrm{d}H}{\mathrm{d}z}\right)-\frac{1}{4}}$$

where H is the scale height of the atmosphere at height z, γ is the ratio of specific heats at constant pressure and constant volume, and h_{mn} is the equivalent depth of the tidal mode In Table 1 average values of h_{mn} and $\lambda_{z,mn}$ are quoted for various semidiurnal modes at a height of 110 km, corresponding to the MSIS83 model of the atmosphere[6] for the five days

In Fig 4 the value of λ_z is used to predict the rate of descent of the phase front for each mode The average rate of descent of the phase front actually observed for $v_{\|}$ is also plotted

The pattern descends from 120 to 100 km at an average rate of 2 5 km h^{-1} suggesting a vertical wavelength of about 30 km. We can therefore eliminate the (2,2) and (2,3) modes, but allowing for uncertainties in the atmospheric model, especially

Table 1 Predicted vertical wavelength

Mode (m,n)	h_{mn} (km)	$\langle\lambda_{z,mn}\rangle$(km)
(2, 2)	7 85	96
(2, 3)	3 67	53
(2, 4)	2 11	38
(2, 5)	1 37	30
(2, 6)	0 96	25

Table 2 Observed vertical wavelength

Date	North-south 101 km	North-south 110 km	North-south 119 km	East-west 101 km	East-west 110 km	East-west 119 km
16 April	—	6.7	—	—	9.0	6.5
14 May	—	6.4	4.7	10.9	8.8	4.6
21 May	8.5	6.3	—	11.8	8.8	6.2
5 November	10.2	6.8	—	11.9	9.6	—
Average	9.4	6.6	4.7	11.5	9.0	5.8
Wavelength (km)		44.6			37.5	

Time of maximum (UT)

in the estimate of dH/dz, and for the perturbing effects of electric fields the observations are compatible with (2,4), (2,5) or (2,6).

More precise evidence involves measurements of the full vector of ion velocity whenever tristatic data are available. In the E-region v_p is largely determined by (1) the semidiurnal tide in the neutral atmosphere, (2) the diurnal tide, and (3) the effect of ionospheric electric fields. At 101 km the effects of the electric field are very small and the amplitude and phase of the semidiurnal and diurnal modes can be determined by a least-squares fit. At greater heights, as the ion-neutral collision frequency decreases, the ion velocity is increasingly affected by the electric field. At 110 and 119 km, however, it is possible to make an accurate correction for this and the phase of the semidiurnal tide can be determined. Table 2 summarizes results for all days with adequate data, and Fig. 4 shows the average time of minimum southward component at each height.

The variation of phase with height corresponds to a vertical wavelength of 41 ± 2 km, which strongly suggests the dominance of the (2,4) mode at high latitudes.

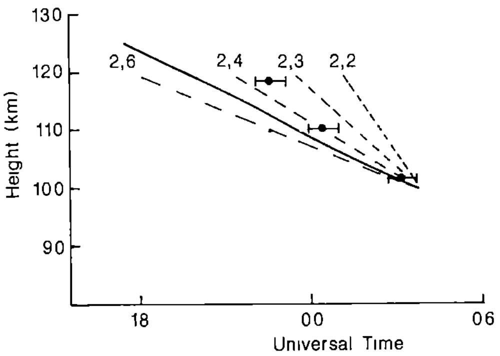


Fig. 4 The predicted rate of descent of several semidiurnal modes (broken lines) and the observed rate-of-descent of the minima of the ion velocity parallel to the magnetic field line (continuous line) and the time of minimum southward horizontal velocity at 101, 110 and 119 km (points with error bars).

We thank the Director and staff of EISCAT for making the observations and analysing the data. EISCAT is supported by the research councils of France, Finland, Germany, Norway, Sweden and the UK.

Received 4 July, accepted 1 September 1986.

1. Alcayde, D. & Bernard, R. *J. atmos. terr. Phys.* **44**, 95-109 (1982).
2. Folkestad, K., Hagfors, T. & Westerland, S. *Radio Sci.* **18**, 867-879 (1983).
3. Rishbeth, H. & Williams, P. J. S. *Q. Jl R. astr. Soc.* **26**, 478-512 (1985).
4. Wilkes, M. V. *Oscillations of the Earth's Atmosphere* (Cambridge University Press, 1949).
5. Chapman, S. & Lindzen, R. S. *Atmospheric Tides* (Reidel, Dordrecht, 1970).
6. Hedin, A. E. *J. geophys. Res.* **88**, 10,170-10,188 (1983).
7. Kirkwood, S. C. *J. atmos. terr. Phys.* **48**, 817-826 (1986).

Long-range transport and deposition of acidic nitrogen species in north-west Europe

Richard G. Derwent* & Katrin Nodop†

* Modelling and Assessments Group, Environmental and Medical Sciences Division, Harwell Laboratory, Oxfordshire OX11 0RA, UK
† EMEP Chemical Coordinating Centre, Norwegian Institute for Air Research, Lillestrøm, Norway

The mechanisms of the long-range transport and deposition of acidic nitrogen compounds over north-west Europe are not yet well understood. Aerosol nitrate concentrations have been steadily rising in southern England[1] over the past three decades, as have the nitrate concentrations in precipitation at almost all sites in north-west Europe, including the most remote[2,3]. The life cycle of acidic nitrogen compounds potentially involves a wide range of trace constituents including nitric oxide (NO), nitrogen dioxide (NO_2), gaseous nitric acid (HNO_3), higher nitrogen oxides (such as N_2O_5), nitrate aerosol and nitrate in precipitation. Here we show that a simple model, incorporating an assumed life cycle of nitrogen-containing trace gases and the known NO sources, can account for much of the nitrate in rain observed over north-west Europe.

The spacial distributions of dry and wet deposition of acidic nitrogen compounds were calculated using a simple trajectory model, the details of which are described elsewhere[4]. The formulation of the model is based on the EMEP (European Monitoring and Evaluation Programme) long-range sulphur transport model[5], although we have made some major simplifications. These have involved the assumption of straight line trajectories, mixing heights held constant at 1,000 m, windspeeds at 8 m s^{-1}, and statistically averaged precipitation-removal rates constant throughout Europe. Figure 1 shows the life cycle of the acidic oxidized nitrogen compounds that has been used and the rate coefficients used to link each of the modelled species. The model consists of a set of coupled differential equations, one for each species including ozone (O_3), representing mass conservation and was solved using FACSIMILE, a variable-order Gear's method program[6] on an IBM PC XT microcomputer.

The major source of acidic nitrogen compounds in north-west Europe is the emission of NO from motor vehicles and stationary combustion. The calculations described here used the NO emissions[7] on the 150 km × 150 km EMEP grid for 1978 from

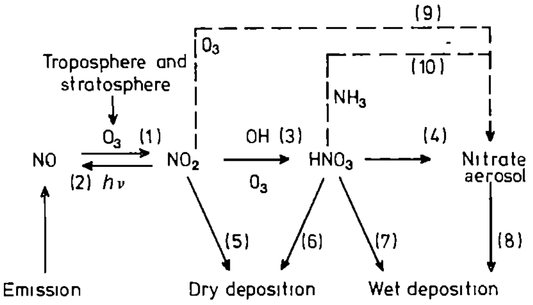


Fig. 1 The life cycle of the nitrogen-containing trace gases and the rate coefficients for the interconversion processes between them (1) 3.8×10^{-4} p.p.b.$^{-1}$ s^{-1}, (2) 2.1×10^{-3} s^{-1}, (3) 3.4×10^{-5} s^{-1}, (4) 3.0×10^{-5} s^{-1}, (5) 4.0×10^{-6} s^{-1}, (6) 2.0×10^{-5} s^{-1}, (7) 9.0×10^{-6} s^{-1}, (8) 1.3×10^{-5} s^{-1}, (9) 1.0×10^{-6} p.p.b.$^{-1}$ s^{-1}, (10) 2.4×10^{-4} p.p.b.$^{-1}$ s^{-1}. The solid lines refer to processes in the base model and the dotted lines to those in the base model + NH_3 + night-time chemistry.

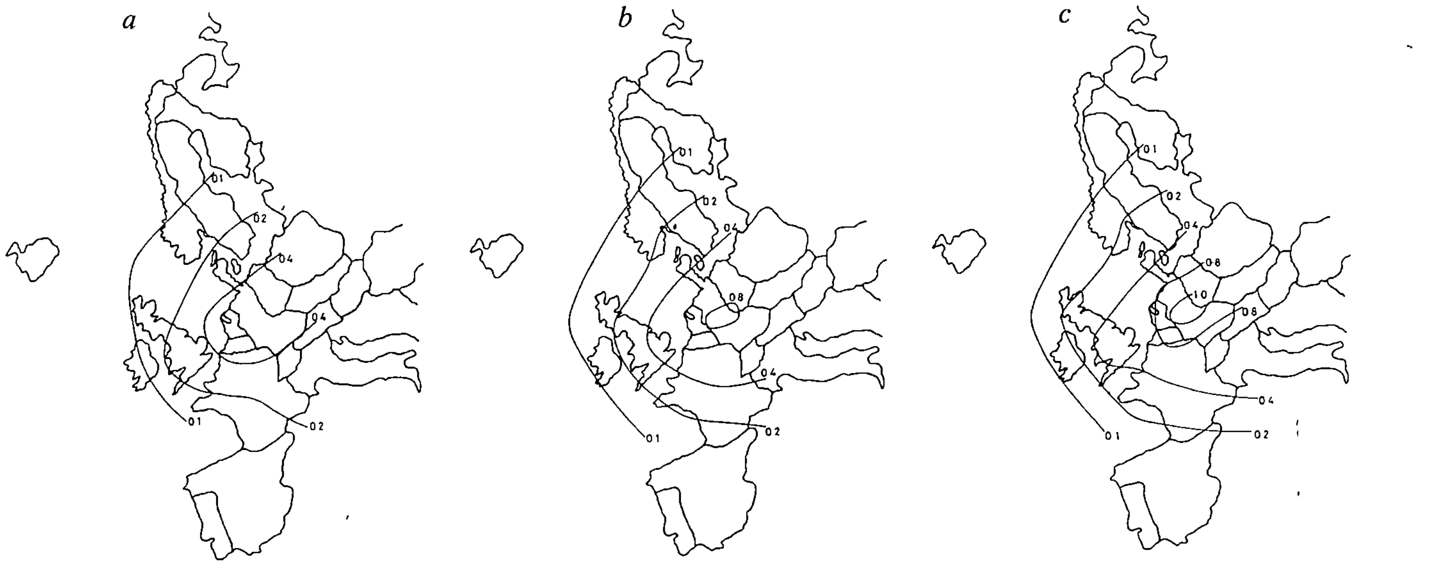


Fig 2 The calculated nitrate concentration in rain over north-west Europe in mg N l^{-1} NO_3 in *a*, the base model, *b*, the base model + NH_3 chemistry, *c*, the base model + NH_3 + night-time chemistry

stationary combustion alone The contribution from motor vehicles to total NO emissions has yet to be specifically treated on the EMEP grid but it is known to be significant overall and could therefore influence both the magnitude and spatial distribution of the emissions The model allows for the oxidation of NO to NO_2 by the well known ozone reaction and then to gaseous nitric acid by reaction with OH radicals A tropospheric background ozone concentration of 30 p p b (30 parts per 10^9, 60 $\mu g\ m^{-3}$) has been employed at the start of each trajectory and is replenished by vertical exchange along the trajectory The OH radical concentration was taken to be that same concentration (2×10^6 molecule cm^{-3}) which would account for the SO_2 oxidation in the EMEP sulphur transport model[5] In the assumed life cycle, HNO_3 is subsequently scavenged by alkaline soil particles of the coarse atmospheric aerosol and by sea-salt aerosol to form nitrate aerosol Dry deposition arises through removal of NO_2 and HNO_3 at the Earth's surface and wet deposition through wash-out of HNO_3 and rain-out of nitrate aerosol Wet deposition was estimated using a constant scavenging coefficient[4] (Fig 1) and an annual rainfall of 1,000 mm None of the processes in Fig 1 is yet adequately defined on a European scale, including the emissions, for oxidized nitrogen species

The concentrations of NO, NO_2, HNO_3 and nitrate aerosol together with the dry and wet deposition rates of the oxidized nitrogen compounds at a given location were obtained by solving the differential equations describing their production and loss for 24 separate 96-h trajectories through to that location, three trajectories within each of eight wind direction sectors 45° wide The long-term mean behaviour at this location was obtained by weighting the results for each trajectory by the frequency of the 96-h back trajectories assigned to each 45° sector in the daily analysis completed by the Meteorological Synthesizing Centre-W for the 636-day period from 1 February 1981 to 29 October 1982 (ref 8) This process was repeated for each of the 64 UN ECE EMEP receptor sites located within the study region for which the above daily sector analysis was available

The calculated concentrations of nitrate in precipitation at each of the 64 receptor sites were plotted out and contours were drawn, joining places of identical rainfall composition (see Fig 2*a*) Most of north-west Europe lay within the 0 1 mg N l^{-1} NO_3 contour under the base model assumptions The total area integrated dry and wet deposition of oxidized nitrogen compounds to the 2×10^6 km^2 land area of Belgium, Denmark, France, FRG, Ireland, Netherlands, Norway, Sweden and the UK amounted to 420 and 450×10^9 g N yr^{-1} The mean residence time for oxidized nitrogen in this base model was found to be 1 2 days, corresponding to a mean transport distance of 800 km The calculated nitrate concentrations in precipitation in Fig 2*a* showed a systematic underestimation of the observed[9] concentration data in Fig 3

In formulating the base model, two additional processes have been specifically omitted (1) scavenging of gaseous nitric acid by ammonia and its subsequent condensation to form ammonium nitrate aerosol, (2) reaction of NO_2 with ozone at night-time to form N_2O_5 and its subsequent reaction with water droplets to form nitrate aerosol We now illustrate the potential importance of these two processes on the spatial distribution of wet nitrate deposition

The formation of ammonium nitrate aerosol was treated by including an additional differential equation for ammonia, representing its emission from animal and industrial sources on the same 150×150 km grid[10], its dry deposition and reaction with HNO_3 to form ammonium nitrate aerosol and its subsequent wet deposition The modelling problem was simplified by assuming no volatilization and redissociation of ammonium nitrate aerosol and by neglecting any competitive scavenging of NH_3 by sulphuric or hydrochloric acid gases These assumptions should exaggerate the importance of ammonium nitrate aerosol formation on the nitrogen life cycle The rate coefficients for the various processes were estimated as few of the relevant processes have yet been studied in any detail except for the thermodynamics of ammonium nitrate formation[11]

The formation of N_2O_5 at night-time was determined using the known rate coefficient of the NO_2+O_3 reaction[12] The NO_3 produced was assumed to be scavenged to form N_2O_5 which, in turn, reacted heterogeneously with water droplets to form nitrate aerosol Constant night-time conditions were taken so that our assumptions should exaggerate the importance of this route in parallel with the optimistic assumptions also made for the ammonium nitrate aerosol formation route The assumptions made for the two additional routes, distinguished from the base model in Fig 1, should therefore be seen to be optimistic but not unreasonably so

Figure 2*b*,*c* shows the nitrate concentrations in rain calculated for the base model + NH_3 and the base model + NH_3 + night-time chemistry A slight widening of the area covered by the 0 1 mg N l^{-1} NO_3 contour is apparent compared with the base model

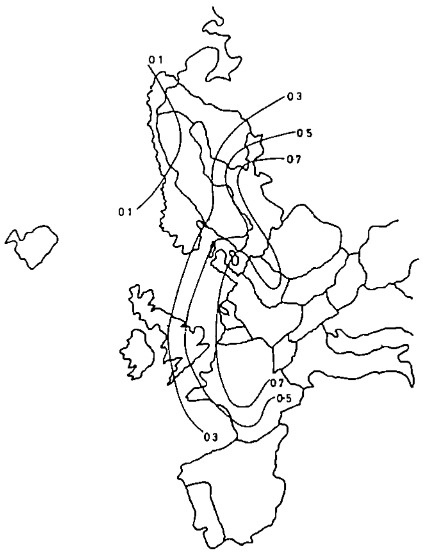


Fig 3 The observed[8] nitrate concentration in rain over north-west Europe in mg N l^{-1} NO_3

results in Fig 2*a*. However, the main change is the marked increase in the magnitude of the peak nitrate concentration in precipitation from 0.64 to 0.82 and 1.08 mg N l^{-1} NO_3. The location of the peak nitrate concentration with precipitation remains in the centre of the FRG with a band of high concentration extending towards the UK through the low countries. A marked increase in the calculated nitrate content of precipitation is noted throughout all the region studied, completely removing the previously noted systematic underestimation of the observations in Fig 3. All these remarks concern the wet deposition of oxidized nitrogen compounds and not the total wet nitrogen deposition which would include a further contribution specifically from the deposition of ammonia and its derivatives.

The effect of the ammonia and night-time chemistry has been to reduce the total area integrated dry deposition of oxidized nitrogen compounds to 140×10^9 g N yr^{-1} and increase their wet deposition to 710×10^9 g N yr^{-1}. The total wet and dry deposition has, however, remained unchanged and all that has been affected is a change in the deposition route from dry deposition to wet deposition. The average residence time of oxidized nitrogen has increased to 1.3 days. Total wet and dry deposition of oxidized nitrogen compounds is controlled by the NO emission strength which has been left unchanged, but the detailed chemistry controls the split between wet and dry deposition. This split is crucial because only wet deposition is directly measurable over north-west Europe. Furthermore, wet and dry deposition may have different ecological effects.

The model results were compared with measurements at 25 EMEP sites for which nitrate in precipitation data are available for either 1981 or 1982 or both. Table 1 shows how the model results correlate with the observations revealing a high degree of correlation ($R^2 \simeq 0.8$) between the model calculations and observations for the 25 sites. The slopes are significantly different between the various model experiments and approach unity only when the ammonia and night-time chemistry is introduced. In all cases, the intercepts are not statistically different from zero indicating the apparent absence of background contribution to nitrate in precipitation. This is an apparent contrast with the analogous situation for sulphate in precipitation[5].

Table 1 Comparison of the observed and model calculated values of nitrate in precipitation at 25 EMEP sites in north-west Europe

Correlations	Slope	Intercept (mg N l^{-1} as NO_3)	Correlation coefficient (R^2)
Base model on data	0.62±0.1	0.02±0.05	0.78
Base model + NH_3 on data	0.80±0.1	0.03±0.07	0.78
Base model + NH_3 + night-time chemistry on data	1.03±0.2	0.01±0.09	0.76

The simple model, with the assumed life cycle and the known NO sources over north-west Europe are therefore able to account for much of the nitrate in rain observed over north-west Europe. However, an adequate explanation requires a better meteorological description and a more detailed understanding of the atmospheric chemistry of NO_x, ammonia and ozone than has been used here and will require a much improved database for ammonia and ozone across north-west Europe in addition to that for NO, NO_2, HNO_3 and nitrate aerosol. Furthermore, as European NO emissions have almost certainly increased over the past 30 years then total wet and dry deposition will have risen proportionately. However, because of possible increased NH_3 emissions[13] and possible increased background ozone concentrations[14], this work demonstrates that the split between wet and dry deposition may also have changed markedly leading to a more dramatic increase in wet nitrate deposition compared with total deposition. Future trends in wet nitrate deposition may not necessarily be dictated solely by NO_x emissions and by the application of NO_x control technologies in vehicles and large combustion plant unless steps are taken to understand and control future NH_3 and ozone concentrations and their possible changes over north-west Europe.

This work was supported in the UK by the Department of Environment and at the Norwegian Institute for Air Research through the UN ECE EMEP Programme and through a contract from the Umweldbundesamt of the FRG. We acknowledge discussions with John Garland, Frank Leeuw, Oystein Hov and Anton Eliassen.

Received 28 July, accepted 3 October 1986

1 Salmon, L, Atkins, D H F, Fisher, E M R, Healy, C & Law, D V *Sci Total Envir* **9**, 161-200 (1978)
2 Rodhe, H K *Skogs o Lantbr Akad Tidskr Suppl* **14**, 32-36 (1982)
3 Kallend, A S, Marsh, A R W, Pickles, J H & Proctor, M V *Atmos Envir* **17**, 127-137 (1983)
4 Derwent, R G *The Nitrogen Budget for the United Kingdom and North West Europe* (ETSU Rep R 37, Harwell Laboratory, 1986)
5 Eliassen, A *Atmos Envir* **12**, 479-487 (1978)
6 Curtis, A R & Sweetenham, W P *FACSIMILE Release H User s Manual* (AERE Rep R11771, Harwell Laboratory, 1985)
7 Semb, A & Amble, E *Emission of Nitrogen Oxides from Fossil Fuel Combustion in Europe* (NILU Rep No 13/81, Lillestrøm, Norway, 1981)
8 Lemhaus, J, Saltbones, J & Eliassen, A *Deposition Patterns and Transport Sector Analysis for a Four-year-period* (The Norwegian Meteorological Institute Report EMEP/MSC-W 1/85, Oslo, 1985)
9 *Long Range Transport of Air Pollutants Monitor 1984*, 1-134 (National Swedish Environmental Protection Board, Solna, Sweden, 1984)
10 Buijsmann, E, Mass, J F M & Asman, W A H *Ammonia emission in Europe* (Institute for Meteorology and Oceanography, State University of Utrecht IMOU Rep R-85-2, Netherlands, 1985)
11 Bassett, M E & Seinfeld, J H *Atmos Envir* **17**, 2237-2252 (1983)
12 Baulch, D L *et al J Phys Chem Ref Data* **13**, 1259-1375 (1984)
13 ApSimon, H M, Kruse, M & Bell, N *Ammonia Emissions and their Role in Acid Deposition* (Imperial College of Science and Technology, London, 1986)
14 Penkett, S A *Nature* **311**, 14-15 (1984)

Baryonyx, a remarkable new theropod dinosaur

Alan J. Charig & Angela C. Milner

Department of Palaeontology, British Museum (Natural History), Cromwell Road, London SW7 5BD, UK

An extremely large claw bone, some 30 cm long, was found in Wealden (Lower Cretaceous) deposits in a Surrey claypit in January 1983. This led to the discovery the following month of the well-preserved skeleton of a new large theropod dinosaur. Only one other theropod specimen comprising more than a few bones had ever been found in Britain, and that discovery was more than a century ago. Indeed, no large theropod, reasonably complete, had previously been discovered in Lower Cretaceous rocks anywhere in the world. Our study so far suggests that the Surrey dinosaur was a typical large theropod in certain respects, resembling, for example *Allosaurus*[1]. In several other respects, however, it differs sufficiently from all known dinosaurs to merit designation as the representative of a new species, genus and family.

More than three years have elapsed since the skeleton was excavated by the British Museum (Natural History) in May–June 1983. We estimate that 70% of the bones collected have already been prepared from the encasing siltstone concretions, although many require further treatment. The immense amount of work involved suggests that it will be at least another four years before a full account of this important creature can be published; we have therefore decided to publish a preliminary account immediately, bestowing a name on the animal.

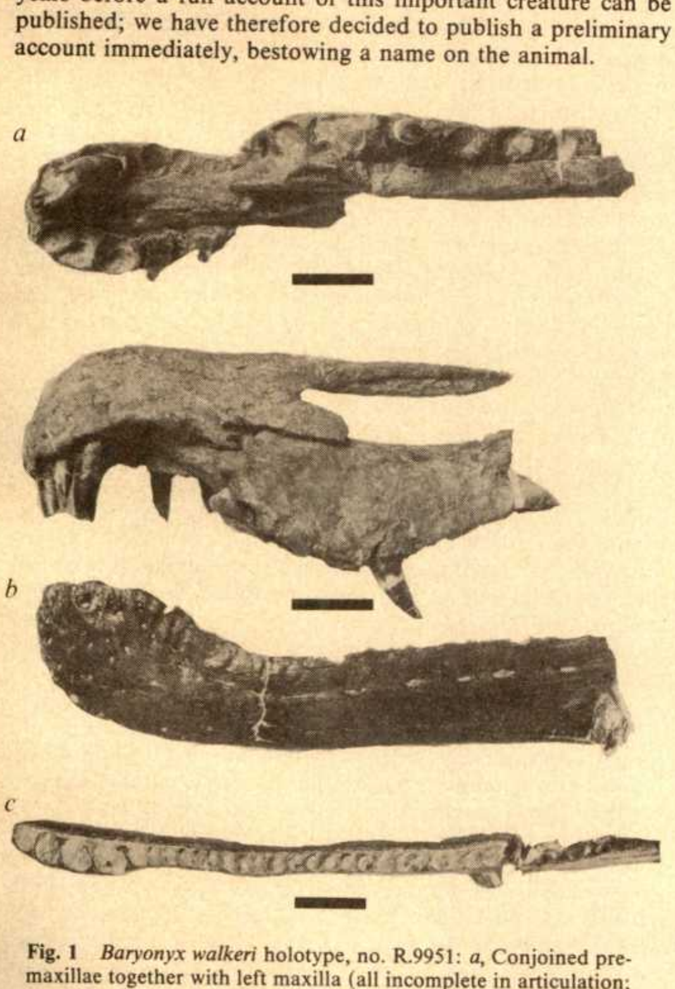


Fig. 1 *Baryonyx walkeri* holotype, no. R.9951: *a*, Conjoined premaxillae together with left maxilla (all incomplete in articulation; in occlusal (ventral) view. *b*, Premaxillae and left maxilla as above, also with incomplete left dentary in approximately correct relative orientation; from the left side. *c*, Left dentary; in occlusal (dorsal) view, to show the full complement of tooth sockets. Scale bars, 5 cm.

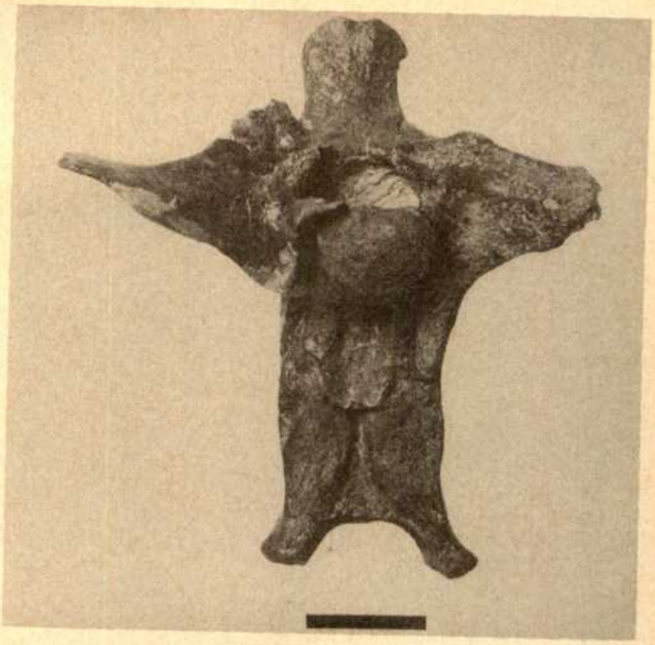


Fig. 2 *Baryonyx walkeri* holotype, no. R. 9951: Occiput, from behind. Scale bar, 5 cm.

Family **Baryonychidae** nov.

Diagnosis. As for genus *Baryonyx*, below.

Genus *Baryonyx* nov.

Derivation of name. Gk, *βᾰρύ*, heavy, strong; *''ονυξ*, talon, claw.

Type-species. *B. walkeri* nov.

Diagnosis. As for species *B. walkeri*, below.

Species *Baryonyx walkeri* nov.

Derivation of name. For Mr William J. Walker, who found the first indication of the skeleton, namely the large claw bone (ungual phalanx).

Holotype. B.M.(N.H.) Palaeo. Dept. R. 9951.

Material. Conjoined premaxillae, front of left maxilla, conjoined nasals, lacrimal and parts of adjacent bones, frontals, anterior end of braincase, occiput, left dentary, some post-dentary bones of lower jaw; teeth in both jaws and in isolation; axis, most of the other cervical vertebrae, some dorsal vertebrae, one caudal vertebra; cervical rib, dorsal ribs, gastralia, chevrons; both scapulae, both coracoids, possible clavicle; both humeri, phalanges of manus including unguals (one of them, not certainly from manus, extremely large); pieces of ilia and pubes, ischium; proximal end of left femur and distal end of right, left fibula lacking distal end; right calcaneum; distal ends of metatarsals, phalanges of pes including unguals; various unidentified pieces; suspected gastroliths.

Locality. Claypit at Smokejacks Brickworks (Ockley Brick Company Limited), Wallis Wood, Ockley, near Dorking, Surrey, England. Map reference TQ13: 113 373.

Horizon. Weald Clay, near the base of the Barremian, Wealden (Lower Cretaceous); approximate age 124 million years.

The bones, though largely disarticulated and scattered, were all found within an area 5 m × 2 m; except where previously disturbed by the bulldozer, they were generally not far from their natural positions, that is most of the pieces of the skull, pectoral girdle and fore-limbs were at one end and most of the pelvic girdle and hind-limbs at the other. They were not crushed or distorted to any significant degree. Many had been broken cleanly before fossilization; most had become encased in sideritic siltstone nodules before the broken ends had weathered or eroded, but a few lay unprotected in clay. So far we have recognized few other vertebrate remains mixed with those of this one individual of *Baryonyx*; these include fish teeth and scales, the humerus of a small *Iguanodon* and a small phalanx. The hardness of the siltstone concretions necessitated the com-

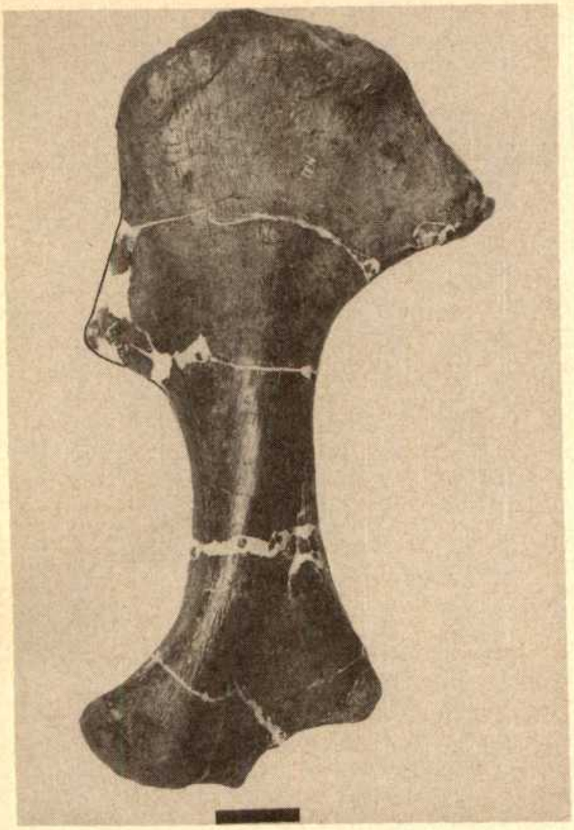


Fig. 3 *Baryonyx walkeri* holotype, no. R.9951: Left humerus, in anterolateral view. Scale bar, 5 cm.

bined use of mechanical (Vacublast, air-abrasive, diamond saws and dental mallets) with chemical (thioglycollic acid[2]) methods of preparation.

Diagnosis. The new theropod is distinguished by a unique combination of characters. These include: the prenarial extension of the snout into an extremely narrow rostrum with a spatulate, horizontal expansion at its end, slightly downturned in lateral view (Fig. 1); the long, low external naris, a flat triangle with its apex directed forwards, situated far back from the front of the snout and ascending gently towards the rear; the seemingly mobile articulation of the premaxilla and maxilla above a subrostral notch (Fig. 1); the small median knob at the posterior end of the conjoined nasals, cruciform in dorsal view, with the anterior limb of the cross drawn forwards into a low thin median crest; the deep occiput (Fig. 2) with paroccipital processes directed horizontally outwards and with basipterygoid processes descending far below the basioccipital and diverging laterally only slightly; the unusually high number of teeth

$$\frac{\mathrm{pm}\ 7+\mathrm{m}?}{\mathrm{d}\ 32}$$

contrast with the usual theropod count of

$$\frac{\mathrm{pm}\ 5+\mathrm{m}\ 16}{\mathrm{d}\ 16}$$

the largest being the third in the premaxilla and the third and fourth in the dentary, the crowns lightly fluted on the lingual side, with very fine serrations (about 7 per millimetre) on the anterior and posterior keels, the roots long and slender; no obvious interdental plates (unlike *Megalosaurus*, where they are very conspicuous, high and separate from each other); the long, strongly opisthocoelous cervical vertebrae with short neural spines, with well developed epipophyses, and with ends of the centra not 'offset' so that when they are correctly apposed the neck does not show the typical theropod upward curve; the relatively well-developed humerus (Fig. 3), with both ends broadly expanded but extremely flattened, offset 35° against each other, shaft massive and almost straight; the ilium with a long, straight posterior process; and at least one pair of huge talons (probably on the hands, Fig. 4). There are other distinguishing characters of a more esoteric nature.

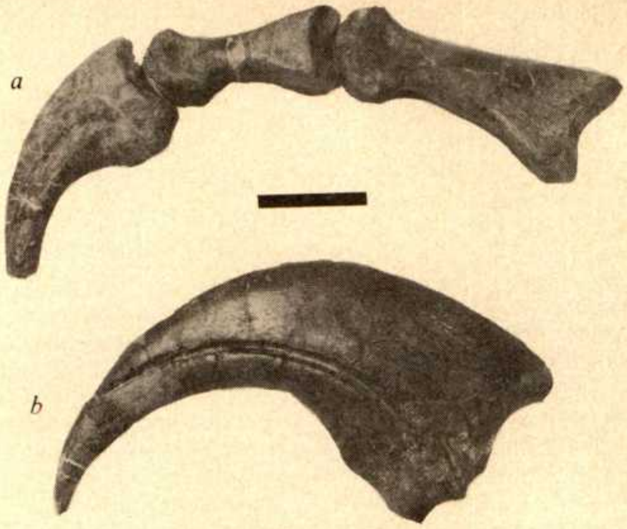


Fig. 4 *Baryonyx walkeri* holotype, no. R.9951: *a*, 'Normal' digit, presumably of hand. *b*, Especially large ungual phalanx. Length of scale bar 5 cm.

Baryonyx shows a strange mixture of primitive and specialized characters. The downturned tip to the snout, the apparent mobility between premaxilla and maxilla and the subrostral notch, the elongation of the cervical vertebrae and the form of the posterior process to the ilium are typical of many thecodontians; the humerus is also primitive in some respects. It is remarkable that so many characters frequently found in Triassic archosaurs should still be present in those from the Early Cretaceous. In contrast, the opisthocoely of the cervical vertebrae and their epipophyses, the lack of a supinator process and ectepicondylar groove to the humerus, the inturned head to the femur and the apparent absence of dermal armour clearly demonstrate that *Baryonyx* is a theropod (as currently defined) with the 'fully improved' stance and gait that distinguish dinosaurs from their thecodontian ancestors. It is not a crocodilian, for neither the skull, the cervical centra, the pelvis nor the ankle shows any of the appropriate specializations. Even so, if the animal is to be considered a theropod, it possesses unique specializations of its own: the terminal 'rosette' to the upper jaw, the extreme narrowness of the skull just behind it, the median knob and crest on the nasals, the large number of teeth (double that of a typical theropod), the lack of an upward curvature to the neck, and the huge offensive/defensive talon. It might also have been quadrupedal to some extent (see below), in which case such stance and gait—whether primitive or secondarily derived—would certainly be unique among theropods.

Therefore we propose that the most proper systematic position for *Baryonyx*, on the evidence presently available, is in a new family (Baryonychidae) of the Theropoda; the information currently available does not indicate a particular relationship to any other group of theropods. (The nature of the articulation between premaxilla and maxilla is somewhat reminiscent of *Dilophosaurus*[3,4] from the Lower Jurassic Kayenta Formation of Arizona, placed by Welles[4] in the Halticosauridae, but *Dilophosaurus* has paired crests on top of its head, whereas *Baryonyx* appears to have only one.)

The only other material that may be assigned to the same family are two fragmentary snouts from the Aptian of Niger, described[5] as the mandibular symphyses of a spinosaurid. They are, however, almost identical to the conjoined premaxillae of

Baryonyx, and comparison with *Spinosaurus*[6] does not suggest spinosaurid affinities.

The possible life habits of *Baryonyx* are unclear. The elongated skull, the long neck and the form and strong development of the humerus might all suggest quadrupedality, facultative at least, even though the flattened slender ends of the humerus makes it unlikely that the stance and gait were predominantly quadrupedal. The lateral position of the external nares, far back from the tip of the snout, seems to preclude an aquatic or amphibious existence, and the form of the postcranial skeleton is wholly suggestive of life on land. Nevertheless the elongated snout with its terminal expansion and the large number of finely serrated teeth may well be indicative of ichthyophagy. Taquet's article[6] on the Niger theropod made the same suggestion, comparing the elongated snouts with those of the Recent *Gavialis gangeticus* and commenting that "*il est tentant d'imaginer des Dinosaures Spinosauridés pêchant le long des fleuves ou au bord des lacs à la manière des Hérons ou des Cigognes.*" We independently conceived a similar idea for *Baryonyx*, though as a quadrupedal predator crouching on the bank rather than as a biped stalking through the shallows. A scavenging habit has also been suggested for the new dinosaur, but the evidence for either mode of life is inconclusive. The associated fossil fauna is dominated by the remains of the herbivorous ornithopod dinosaur *Iguanodon* and of many insects[7], the latter under continuing study by Jarzembowski.

Whether the huge talon was one of an only pair remains unknown; likewise whether it was borne by the fore-limb or by the hind. The ungual is exactly like those of *Allosaurus*, only much bigger. Clearly it was a weapon of offence and/or defence; but it seems likely that the great size of *Baryonyx* (it may have weighed a couple of tons or more) would have precluded any possibility of emulating its contemporary *Deinonychus* from the Cloverly Formation of Montana[8], a biped of much smaller dimensions which bore a large, laterally-compressed, sickle-like claw on the second digit of the pes and which balanced on the other hind-leg while slashing at its victim. If, however, the great talon of *Baryonyx* was on the manus, it could be connected with the powerful development of the humerus. All this must have some bearing on facultative quadrupedality (the claw would have had to be held clear of the ground while the animal was walking) and on ichthyophagy; could *Baryonyx* have used the claw, like a grizzly bear, for 'gaffing' fish out of the water? *Baryonyx* is the dinosaur that in 1983 received world-wide publicity, and became nicknamed 'Claws'. We are expecting its remains to provide a great deal more information over the next few years.

Our thanks are due to the many people who helped in different ways, too numerous to specify. They include Mr William Walker, the Directors and staff of the Ockley Brick Company Limited, the collecting teams (Museum staff and others), the Museum photographers, Mr Philip Palmer, and Mr Ronald Croucher and his colleagues in the Museum's Palaeontological Laboratory.

Received 4 August; accepted 1 October 1986.

1. Madsen, J. H. Jr *Bull. Utah geol. miner. Surv.* **109**, i-xii, 1-163 (1976).
2. Howie, F. M. P. *Curator* **17**, 159-166 (1974).
3. Welles, S. P. *Bull. geol. Soc. Am.* **65**, 591-598 (1954).
4. Welles, S. P. *Palaeontographica*, Abt. A**185**, 85-180 (1984).
5. Taquet, P. *C.r. hebd. Séanc. Acad. Sci., Paris* **299**, 217-222 (1984).
6. Stromer, E. *Abh. bayer. Akad. Wiss.* **28**, 1-32 (1915).
7. Jarzembowski, E. A. *Proc. Geol. Ass.* **87**, 443-446 (1977).
8. Ostrom, J. H. *Bull. Peabody Mus. nat. Hist.* **30**, 1-165 (1969).

Functional architecture of cortex revealed by optical imaging of intrinsic signals

Amiram Grinvald*†‡, Edmund Lieke*, Ron D. Frostig*, Charles D. Gilbert* & Torsten N. Wiesel*

* Laboratory of Neurobiology, Rockefeller University,
1230 York Avenue, New York, New York 10021, USA
† Department of Neurobiology, Weizmann Institute of Science,
Rehovot 76100, Israel
‡ T. J. Watson Research Center, IBM, Yorktown Heights,
New York 10598, USA

Optical imaging of cortical activity offers several advantages over conventional electrophysiological and anatomical techniques. One can map a relatively large region, obtain successive maps to different stimuli in the same cortical area and follow variations in response over time. In the intact mammalian brain this imaging has been accomplished with the aid of voltage sensitive dyes[1-5]. However, it has been known for many years that some intrinsic changes in the optical properties of the tissue are dependent on electrical or metabolic activity[6-13]. Here we show that these changes can be used to study the functional architecture of cortex. Optical maps of whisker barrels in the rat and the orientation columns in the cat visual cortex, obtained by reflection measurements of the intrinsic signal, were confirmed with voltage sensitive dyes or by electrophysiological recordings. In addition, we describe an intrinsic signal originating from small arteries which can be used to investigate the communication between local neuronal activity and the microvasculature. One advantage of the method is that it is non-invasive and does not require dyes, a clear benefit for clinical applications.

The imaging of changes in membrane potential in central nervous system (CNS) preparations with voltage sensitive dyes, which are usually associated with fast signals, is often distorted by much slower intrinsic signals[6,11]. The observation of only slow signals from the surface of monkey striate cortex stained with a voltage sensitive dye[14] prompted us to test whether slow reflectance signals could be observed without using any dye, and to compare the relative merits of intrinsic and dye-related imaging. For optical recording from the rat somatosensory cortex, the exposed cortex was viewed with a compound microscope equipped with fluorescence epi-illumination through a glass-coverslip sealed chamber. The rat somatosensory cortex includes a region that is subdivided into smaller areas, known as barrels, each of which responds to stimulation of a single mystacial whisker[15]. Using the photodetector array[16] to monitor changes in reflected light from an active barrel field, we observed a decrease in reflectance in response to whisker stimulation. In dye stained preparations, this intrinsic signal was distinguished from the dye signal by measurement at a wavelength range outside the absorption band of the dye[6]. This made it possible to compare directly maps obtained by the reflected light and voltage sensitive dye signals in response to stimulation of one whisker. Both showed signals in a well-localized cortical region and the centres of the active areas overlapped very closely (Fig. 1*a*, *b*). Although the reflected light signal was much slower than the dye signal (Fig. 2*a*), representing different processes, it was still as useful for visualizing the active cortical region.

To determine the optimal conditions for the measurement and to characterize the signal source, we examined the time course and wavelength dependence of the signal. The reflected light signal in the rat somatosensory and in the cat and monkey visual cortex had a slower rise time (>1 s) than the dye related signals (Fig. 2*a*, *b*). This time course is clearly very different from neuronal electrical activity. By comparing the optical signals in response to long and short stimuli we found that the rise time and time to peak were nearly independent of stimulus duration (Fig. 2*d*). Examination of the relaxation time course of the signal was important for choosing the appropriate inter-stimulus interval. Short duration stimuli (0.5-2 seconds) pro-

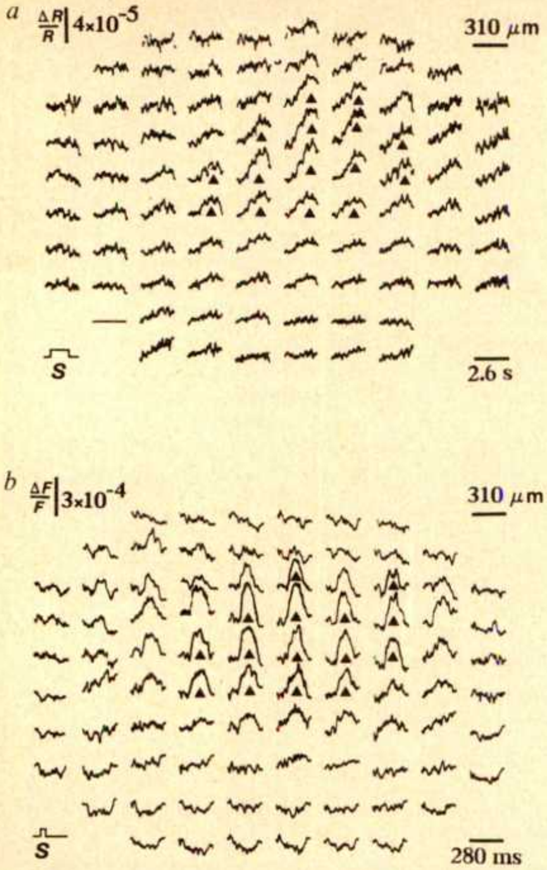


Fig. 1 Mapping of a single barrel in rat cortex with the intrinsic signal and its confirmation with a voltage-sensitive dye. *a*, The spatial distribution of the reflected light signals in response to a mechanical stimulation of a single whisker. Simultaneous intensity changes in reflected light were measured with 96 photodetectors, averaged over 200 trials. Signals marked by triangles are the largest 15 signals selected by computer integration. A wavelength of 665 to 750 nm was used. A semi-silvered mirror was used as a reflector. *b*, As *a* but using fluorescence mapping, averaged over 80 trials. Again, the largest 15 signals were marked by triangles. The centres of the two maps of the barrel are superimposed. Note that in this experiment the area of activation with the intrinsic signal is smaller than that seen with the dye signal. (We do not know if the difference is significant because the two measurements were not simultaneous and the fluorescence measurement was done first.) The cortex was stained with 0.1 mg ml^{-1} RH-795 for 90 min. We used an excitation filter at 530 nm (bandwidth 70 nm), an FT-580 dichroic mirror and an RG610 barrier filter. Common mode noise due to the heartbeat artefact and respiratory movements was removed by subtracting a signal-free trace from the rest of the traces. The duration of the stimulus (*S*), the size of cortical area viewed by each detector, the time scale, and the fractional change in light intensity ($\Delta R/R$ or $\Delta F/F$) reaching the detectors, are marked at the corners of the map. Note the difference in time scale of *a* and *b*. For a discussion of the area of an optically detected barrel relative to its anatomically determined size see the discussion in ref. 5. In similar mapping experiments, using the intrinsic signal, two individual barrels could be resolved if two whiskers were stimulated.

Methods. The optical monitor is described elsewhere[3,6,16]. The detector amplifiers have a sample-and-hold option that facilitated the high gain d.c. recordings. To minimize the large noise associated with heartbeat, the stimulus and data collection were triggered by the heartbeat, and each stimulus associated trace was subtracted from a trace obtained without stimulus[5,17]. We used 22 rats, 21 cats and one monkey. Rats were anaesthetized with an intraperitoneal injection of urethane. Cats and monkeys were anaesthetized with continuous infusion of sodium pentothal, paralysed with succinylcholine or curare/gallamine, and artificially respirated. The visual stimulus was a square-wave grating, 0.3–1 cycles per degree, drifting at a rate of 2–3 cycles per s. The surface of the cortex was positioned perpendicular to the axis of the microscope objective by a modified stereotaxic apparatus. Most of the light reaching the photodetectors was stray light (80–90%) reflected from the epi-illuminator attachment, the objective lenses, and the coverglass of the chamber attached to the skull. Note that the real fractional change of reflected light from cortex is therefore 5–10-fold larger than that shown on the vertical calibration scales. No attempt was made to minimize the stray reflections because they did not contribute to the noise originating from movements of the cortical surface.

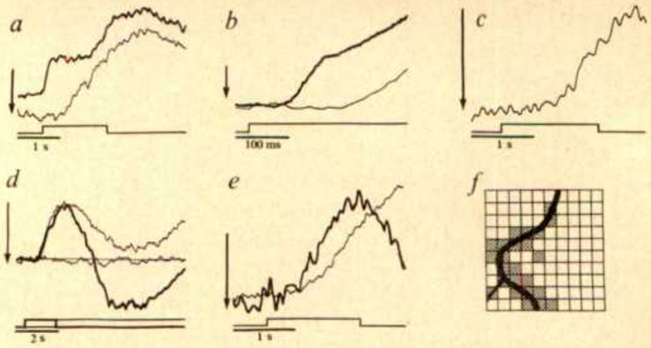


Fig. 2 Characterization of the reflected-light signal in rat, cat and monkey cortex. *a*, Comparison between reflected light signal (fine trace) and the fluorescence signal (bold trace) in the rat somatosensory cortex. Cortical activity was evoked by a 15° deflection of 6 whiskers. The two signals were normalized to the peak. *b*, Comparison between reflected light signal (thin trace) and signal from voltage-sensitive dyes (bold trace), measured from the same cortical area of the cat visual cortex (note the expanded time scale). The amplitudes were normalized so that the late component of the fluorescence would have a similar slope to the light reflection signal. Note the two components of the fluorescence signal in *a* and *b*. (The late component of the fluorescence signal is produced by the intrinsic signal, and not by a change in membrane potential. Cortex stained with non-voltage-sensitive fluorescent dyes, such as fluorescein, gave signals with a time course identical to the intrinsic signal.) *c*, The reflected light signal from the monkey visual cortex shown on a slow time scale. *d*, The relaxation time course of the signal. The shape of the intrinsic signal depended on the duration of the stimulus, having an undershoot to a short (1.5 s) stimulus, but not to a longer (8 s) stimulus. The flat trace was obtained without a stimulus to show the noise associated with the measurements. *e*, Comparison of the time course of the reflected light signal from the barrel (bold line) in the rat cortex and the arterial signal outside the barrel field (fine line). The arterial signals were seen along the full course of an artery (3 mm) which traversed the stimulated barrel. *a–e*, Vertical arrows, size and direction of fractional change in reflected light, $\Delta R/R$ (fine traces; scale in *a–c*, 10^{-4}; *e*, 2×10^{-4} bold trace: scale in *e*, 10^{-4}; scale in both traces *d*, 5×10^{-4}) or of fractional change in fluorescence, $\Delta F/F$ (bold traces; scale in *a*, 10^{-3}; scale in *b*, 2×10^{-4}). *f*, The spatial distribution of the largest reflected light signals monitored at a wavelength of 850 nm (bandwidth of 40 nm), in cat cortex. Detectors that had a signal larger than 50% of the maximal signal are shaded in grey. Their distribution co-localized with an artery (150 μm in diameter) in the field of view even at this wavelength.

duced a large and slow (10 s) undershoot following the peak of the signal. Because of the slow recovery time of the undershoot, to avoid distortion of the signals it is necessary to use interstimulus intervals of at least 10 s for visual stimuli lasting 1–2 s.

The cerebral vasculature is the source of one component of the intrinsic signal. In the rat experiment the largest signals were observed along arteries, but not veins, and only when the arteries passed through an activated barrel. The signal associated with arteries was delayed relative to the signals from the active barrel (Fig. 2*e*). The arterial signal was seen at wavelengths ranging from 400 to 1,100 nm (for example, 850 nm, Fig. 2*f*). The other signal components that were useful for functional mapping showed the optimum signal-to-noise ratio at a wavelength range of 610–750 nm. We observed that when using another dye (RH414, ref. 17) the arteries constricted, and only the artery-associated signal disappeared, suggesting that the signal represented changes in blood flow in response to local neuronal activity[18,19]. This arterial signal could undoubtedly be a useful tool in studying the relationship between neuronal activity and microcirculation but when making functional maps it can introduce distortions. The distortions can be eliminated from the maps with the appropriate analysis, as shown below.

Having optimized the stimulus and recording conditions we obtained optical maps for different stimulus orientations in the

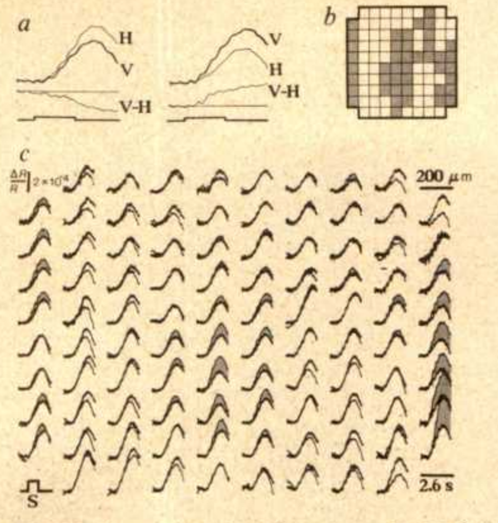


Fig. 3 Visualization of the orientation columns in cat visual cortex. *a*, Demonstration of the signal sensitivity to the orientation of the stimulus. Right traces: the response to vertical drifting gratings (V) and to horizontal drifting gratings (H). The difference between these two traces is shown in the third trace ($V-H$). An electrode penetration at this site recorded from cells which responded only to vertical gratings. Left traces: as for right, but the optical signals are from detectors that monitored cortical area in which units responded only to horizontal stimuli. *b*, A two-state map of the orientation columns. For each pixel, we integrated the signals for the vertical and horizontal stimuli. Shaded detectors, those giving a larger response to horizontal than to vertical stimuli. *c*, The raw data. Two superimposed patterns of reflected light signals in response to horizontal (fine lines) and vertical (bold lines) drifting gratings. These experiments were interlaced. Positions where the response to horizontal stimuli was larger than the response to vertical stimuli are shaded in grey. The pattern of optical signals may seem puzzling as optical signals are seen everywhere, also in areas where single unit responses were not recorded for a given stimulus orientation[20]. We attribute this to a component of the signal that is independent of stimulus orientation and lateral spread of activity along dendritic and axonal processes. The signals were averaged over 100 trials. The signal to noise ratio varied from animal to animal and also depended on the physiological state of the animal. In some experiments the pattern became apparent after averaging 10 trials. In other experiments 200 trials were required. Objective criteria such as excessive noise or unusually slow signals could be used to predict whether an optical map is accurate.

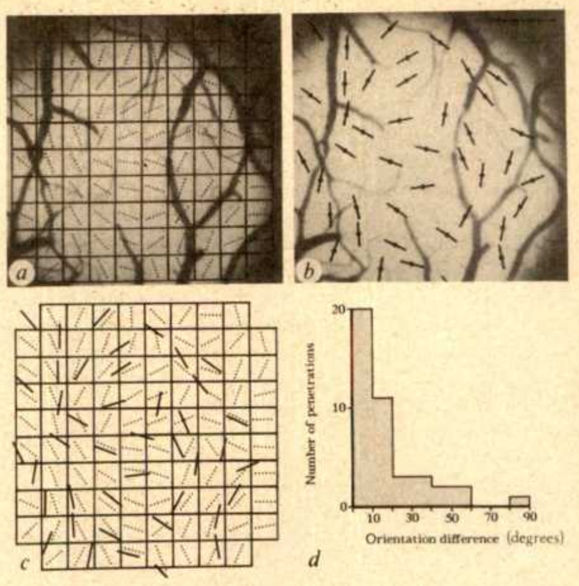


Fig. 4 Electrophysiological confirmation of the optical map of orientation columns. *a*, Optical map of the optimal stimulus orientation, for each cortical site. Since the size of the signal at any given site depends on the inhomogeneous distribution of signal sources as well as the orientation of the stimulus, we took the ratio (rather than the difference) of the integrals of the optical responses to the orthogonal stimulus for each pixel. The ratio approach minimized map distortions produced by inhomogeneous distribution of signal sources (for example, arteries or capillaries). From this ratio we calculated a vector for each pixel. The angle of the optimal stimulus orientation (dotted line) was calculated by vectorial addition of two sets of experiments (see text). The position of the photodiode array (grid) is shown relative to the cortical surface. This map was highly reproducible. A second set of experiments done several hours apart provided an identical orientation map, such that in 95% of the detectors the average difference in the optimal orientation angle determined by the two independent experiments was less than 5°. *b*, The results of the electrophysiological mapping with 42 perpendicular penetrations with tungsten electrodes. The preferred angles of orientation for the single units were determined with a hand-held projector. The location of each penetration was marked on a photograph of the cortical surface. Scale marker, 500 μm. *c*, Superimposed optical (dotted lines) and electrophysiological maps (bold lines). Although there were discrepancies at a few sites, a pixel adjacent to the site of electrode penetration often had an orientation very close to that determined electrophysiologically. Note that large deviations were not associated with the vascular pattern. *d*, Evaluation of the correspondence between the orientations determined electrophysiologically and optically. The histogram shows the deviations between the optimal angle determined by single unit recordings and the optically determined angle at that site or within 100 μm away. Several sources of error are associated with this comparison including the error in the precise marking the sites of the electrode penetrations and the fact that each pixel sums the activity over an area of 200 × 200 μm. Note that 74% of the points fell within 20°, the estimated accuracy of the single unit recordings.

cat visual cortex. These maps were compared to maps generated by multiple electrode penetrations. Although signals to both vertical and horizontal stimulus presentations were seen everywhere, in restricted regions of cortex, the optical signals showed a larger response to a horizontal drifting grating (Fig. 3*a*, left, thin trace) than to a vertical grating (thick trace). Other regions showed the opposite behaviour, giving a larger optical signal to the vertical grating (Fig. 3*a*, right). In other cortical sites the optical signals were nearly identical and in those sites single units did not respond to the pair of stimulus orientations used. The difference between the signals to orthogonal stimuli (vertical (V) − horizontal (H)) showed the signal component that was sensitive to orientation. The spatial distribution of the regions preferring the horizontal stimulus is shown as a two-state coded map in Fig. 3*b*. The raw data are shown in Fig. 3*c*.

We combined the results from vertical and horizontal stimulus presentations (Fig. 3*c*) with results from presentations at orientations of 45° and 135° to obtain a more precise map of the orientation columns. We determined the optimal orientations of each pixel by vector addition (Fig. 4*a*). To confirm the optical map the same area of cortex was mapped, using a double-blind procedure, by making a series of 42 electrode penetrations (Fig. 4*b*). The two maps showed a good correspondence (Fig. 4*c*, *d*). From similar experiments in three different cats we conclude that the intrinsic signals can be used to map the orientation columns.

Our findings from the intrinsic and dye signals leads to several conclusions concerning their relative merits. The principle advantage of optical mapping using voltage-sensitive dyes is that the time resolution is <1 millisecond, while that of the reflected light signals is in the order of seconds. The better time resolution of the dyes is essential for analysing the cortical circuitry and for understanding the sequential flow of information and its processing at different cortical sites. However, at present the dyes have several shortcomings; some of them may induce pharmacological side effects and cause photodynamic damage upon prolonged or intense illumination, they bleach and the depth of penetration into the cortex is limited[1-5]. Often, the dye signals are distorted by the intrinsic signals[6,11].

A voltage-sensitive dye, NK2367, was used by Blasdel and Salama to study the functional architecture of the monkey striate cortex[14]. There are several discrepancies between their work and ours. The first issue is whether their signal came from the dye, as they claim, or from an intrinsic source. The slow time course of the signals seen in both studies suggests that their signal did not come from the dye, as in all other preparations using NK2367 the signal included a fast component[21-24] even if a slow glial signal was also present[25]. Furthermore, the standard evidence to attribute a signal to a dye is the observation of different signals at different wavelengths, and this evidence was not provided. Another issue is the usefulness of the intrinsic signal for mapping. Although Blasdel and Salama observed an intrinsic signal, they did not find it useful for functional mapping. Finally, there is the question of distortion of maps by the cortical vasculature. As we show, the blood vessel signals only appear when the cortex is active (that is, in the presence of a stimulus). The fact that no pattern is seen in the control experiment, in the absence of a visual stimulus leads Blasdel and Salama to conclude that the vascular pattern does not contribute to the maps. In their work, however, the vascular pattern is visible in the functional maps. Although we see arterial signals in our raw data, there is no arterial contribution to the functional maps (for example, Fig. 4), possibly because of the method of analysis and data acquisition, the low spatial resolution of the recording system or a lower noise level.

The exact origin of the multiple components of the reflected light signal will determine the ultimate spatial resolution and remains to be investigated. In addition to the arterial component, which probably reflects changes in blood volume or flow, a second component probably originated from local activity light scattering signals as described in previous reports[6-11,26]. The third component may originate from local changes in the absorption associated with the conversion of oxyhaemoglobin to haemoglobin[27-29] in the capillaries, in response to metabolic demands. In fact we have recently found that this component is a significant constituent of the intrinsic signal.

We thank Drs L. Katz and E. Kaplan for their useful comments. Supported by NIH grants NS14716 and EY05253, NSF grants BNS9318794 and 8351738 and grant AFOSR-85-0301 from the US Air Force and by IBM.

Received 21 May; accepted 24 September 1986.

1. Salzberg, B. M., Davila, H. V. & Cohen, L. B. *Nature* **246,** 508-509 (1973).
2. Salzberg, B. M. in *Current Methods in Cellular Neurobiology* (ed. Barker, J. L.) 139-187 (Wiley, New York, 1983).
3. Grinvald, A. *A. Rev. Neurosci.* **8,** 263-305 (1985).
4. Cohen, L. B. & Lesher, S. L. in *Optical Methods in Cell Physiology* (eds DeWeer, P. & Salzberg, B. M.) 71-99 (Wiley, New York, 1986).
5. Orbach, H. S., Cohen, L. B. & Grinvald, A. *J. Neurosci.* **5,** 1553-1564 (1985).
6. Grinvald, A., Manker, A. & Segal, M. *J. Physiol., Lond.* **333,** 269-291 (1982).
7. Hill, D. K. & Keynes, R. D. *J. Physiol., Lond.* **108,** 278-281 (1949).
8. Cohen, L. B., Keynes, R. D. & Hille, B. *Nature* **218,** 438-441 (1968).
9. Tasaki, I., Watanabe, A., Sandlin, R. & Carnay, L. *Proc. natn. Acad. Sci. U.S.A.* **61,** 883-888 (1968).
10. Lipton, P. *J. Physiol., Lond.* **231,** 365-383 (1973).
11. Salzberg, B. M., Obaid, A. L. & Gainer, H. *J. gen. Physiol.* **86,** 395-411 (1985).
12. Chance, B., Cohen, P., Jobsis, F. & Schoener, B. *Science* **137,** 499-508 (1962).
13. Jobsis, F. F., Keizer, J. H., LaManna, J. C. & Rosental, M. *J. appl. Physiol.* **43,** 858-872 (1977).
14. Blasdel, G. G. & Salama, G. *Nature* **321,** 579-585 (1986).
15. Woolsey, T. A. & Van der Loos, H. *Brain Res.* **17,** 205-242 (1970).
16. Grinvald, A., Cohen, L. B., Lesher, S. & Boyle, M. B. *J. Neurophysiol.* **45,** 829-840 (1981).
17. Grinvald, A., Anglister, L., Freeman, J. A., Hildesheim, R. & Manker, A. *Nature* **308,** 848-850 (1984).
18. Roy, C. S. & Sherrington, C. S. *J. Physiol., Lond.* **11,** 85-108 (1890).
19. Lassen, N. A. & Ingvar, D. H. *Experientia* **17,** 42-43 (1961).
20. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160,** 106-154 (1962).
21. Salzberg, B. M., Grinvald, A., Cohen, L. B., Davila, H. V. & Ross, W. N. *J. Neurophysiol.* **40,** 1281-1291 (1977).
22. Grinvald, A., Salzberg, B. M. & Cohen, L. B. *Nature* **268,** 140-142 (1977).
23. Senseman, D. M. & Salzberg, B. M. *Science* **208,** 1267-1271 (1980).
24. Ross, W. N. & Krauthamer, V. *J. Neurosci.* **4,** 659-672 (1984).
25. Konnerth, A. & Orkand, R. K. *Neurosci. Lett.* **66,** 49-54 (1986).
26. Cohen, L. B. *Physiol. Rev.* **353,** 373-418 (1973).
27. Millikan, G. A. *Proc. R. Soc.* **B123,** 218-241 (1937).
28. Jobsis, F. F. *Science* **198,** 1264-1267 (1977).
29. Mook, G. A., Osypka, P., Strum, R. E. & Wood, E. H. *Cardiovasc. Res.* **2,** 199-209 (1968).

Control of malarial invasion by phosphorylation of the host cell membrane cytoskeleton

K. Rangachari, A. Dluzewski, R. J. M. Wilson* & W. B. Gratzer

Medical Research Council Cell Biophysics Unit, King's College,
Drury Lane, London WC2B 5RL, UK
* National Institute for Medical Research, The Ridgeway,
London NW7 1AA, UK

It has been shown that the entry of the malaria parasite into the red blood cell requires the presence of ATP in the host cell cytoplasm[1,2]. In red blood cell ghosts that contain no ATP the receptor on the extracellular surface remains in place and parasites will bind to the membrane, but will not enter[3]. ATP is thus necessary for one of the steps[4] in the invasion sequence that follows recognition and attachment. The process of entry appears to involve the active participation of the host cell membrane cytoskeleton[5]. We have suggested[2] that the function of the intracellular ATP may be to regulate phosphorylation of the cytoskeleton. We now present evidence that the activity of the membrane-associated cyclic AMP-independent kinase of the red blood cell is inseparable from invasion; the active substrate may be spectrin.

Kinases are found in both the membrane and cytoplasm of the human red cell, and there is extensive turnover of phosphoryl groups on cytoskeletal proteins. Spectrin is phosphorylated in the cell by a cAMP-independent (casein) kinase at four sites, clustered at one end of the smaller (β) subunit[6]. Early evidence that phosphorylation at these sites was linked to metabolic shape control[7] did not survive closer scrutiny[8,9] and there is no known function for the phosphorylation. However, it does appear that the phosphorylation state of spectrin can be perturbed by external stimuli[10,11] and migration of intramembrane particles may also be linked to spectrin phosphorylation[10,12]. Rearrangement of intramembrane particles appears to occur during entry of the malarial parasite. A bare area of membrane develops in the region of contact[13] and a dense zone, consisting almost certainly of clustered intramembrane particles, can be discerned in the electron microscope at the parasite-host cell junction when the merozoite begins to move into the cell[4]. To establish whether invasion depends on phosphorylation of spectrin (or of other membrane proteins) we used our earlier observation that malaria parasites can, under defined conditions, be made to enter resealed red blood cell ghosts[14] and the discovery that the cAMP-independent kinase can be extracted from the membrane by a medium of high ionic strength[9].

We found that fresh ghosts hypotonically lysed in the presence of ATP could be subjected to one buffer wash before resealing in the presence of the original haemolysate and 1 mM ATP, without loss of susceptibility to invasion. Freshly lysed ghosts were exposed to a buffer containing 0.65 M potassium chloride to extract the kinase[9], then recovered by centrifugation and resealed with haemolysate and ATP. Inoculation of these ghosts with purified schizonts of *Plasmodium falciparum* revealed drastically reduced susceptibility to invasion (Table 1).

A crude preparation of the cAMP-independent kinase can be made from a high-salt extract of membranes[15]. When this was added (~100 μg per ml final protein) to the haemolysate, invasion of salt-extracted ghosts was substantially restored. This enzyme preparation contains a wide range of membrane proteins, of which the kinase makes up a very small part. It is thus open to question whether it was kinase, rather than some other

Table 1 Effect of cAMP-independent kinase on invasion of red cells by *P. falciparum*

Sample	Invasion ratio* 1	2	3	4	5	6	7	8	Average
Untreated	1.3	3.0	1.5	2.0	2.7	0.9	0.7	—	1.73±0.36
Kinase extracted	0.6	0.4	0.3	0.9	0.8	0.1	0	0.2	0.41±0.12
Kinase replaced	1.9	1.2	1.0	1.8	2.0	0.9	0.8	1.0	1.33±0.19

* Ratio of fraction of parasitized cells after one generation (~24 h) to that in the culture after addition of inoculum. Numbers 1–8 are separate experiments with different cultures: the kinase in experiments 1–5 was the purified enzyme, and in 6–8 was the crude preparation. Averages and standard mean errors are given.

Ghosts were prepared by hypotonic lysis of washed human red cells in 5 mM phosphate, 2 mM magnesium chloride, 2 mM ATP, *p*H 7.4. The membranes were collected by centrifugation and extracted with four volumes of 0.65 M potassium chloride, 0.1 mM EDTA, 0.25 mM dithiothreitol, *p*H 7.4, on ice for 15 min to remove kinase[9]. The pellet was washed once with 0.14 M potassium chloride, 2 mM ATP. Extracted and untreated ghosts were resealed by incubation at 37 °C with a diluted haemolysate (usually 1:8 with isotonic saline, containing 2 mM magnesium chloride, 2 mM ATP) and a kinase preparation as required. The total added protein, estimated colorimetrically[19], was ~100 and $<0.4\ \mu g\ ml^{-1}$, in the case of the crude and purified kinase, respectively. The resealed ghosts were twice washed with culture medium (RPMI 1640 + 10% human serum) and suspended at 50% haematocrit. Schizonts of *P. falciparum* were purified from *in vitro* cultures[17,18] and were added to the ghosts. The extent of parasite invasion was determined after ~24 h by counting the number of new parasites in 1,000 cells in thin smears, stained with Giemsa's solution. A crude kinase preparation was obtained by extraction of cytoskeleton-depleted red cell membranes with a medium containing 0.5 M NaCl as described[15]. The extract was fractionated, after precipitation with ammonium sulphate, by two successive elutions from Sephadex G-100, the first in 0.1 M, the second in 0.5 M NaCl. Two zones of activity eluted from the second column, the first with spectrin in the void volume, the second (free kinase) later. Activity was assayed with [γ-^{32}P]ATP and casein[15]. The peak active fractions were concentrated by vacuum dialysis.

membrane constituent, that was responsible for restoration of capacity for invasion. A further fractionation of the kinase preparation was therefore undertaken using the chromatographic procedure described in ref. 16. The activity elutes at high ionic strength from Sephadex G-100, in two regions, one of high relative molecular mass (M_r), which also contains extracted spectrin, the other of M_r ~34,000. The latter is free kinase, and contains little contaminating protein. The total protein added back in this form to the ghosts before resealing had a concentration of $<0.4\ \mu g\ ml^{-1}$. Phosphorylation of membrane proteins was measured by autoradiography of electrophoretic gels of the total solubilized membrane protein after incubation of hypotonic ghosts with ^{32}P-labelled ATP. The single high-salt extraction (Table 2) removed about half of the endogenous cAMP-independent kinase activity. A second extraction led to extensive vesiculation and misshapen residual ghosts, and was therefore not used. Nevertheless, ghosts in which spectrin phosphorylation was reduced by up to 50% showed consistent reduction of invasion by *P. falciparum* (60–80%; Table 1). Addition of the kinase restored invasion to essentially the original level. Kinase which did not cause phosphorylation (inactived by some days' storage) also did not restore invasion. The requirement for ATP is absolute; non-hydrolysable analogues are ineffective[2].

Table 2 Effect of salt-extraction and replacement of kinase on phosphorylation of spectrin in ghosts

Sample	Relative areas under zone profiles Protein stain	Autoradiograph	Relative ^{32}P incorporation
Untreated	1.00	1.00	1.00
Salt-extracted	1.15	0.60	0.52
Kinase replaced	1.03	1.27	1.23

Unsealed ghosts were incubated with 1 mM [γ-^{32}P]ATP (20–40 c.p.m. $pmole^{-1}$) in hypotonic buffer (as in Table 1) on ice for 20 min. The ghosts were made isotonic and resealed at 37 °C, washed twice in isotonic saline and lysed in the hypotonic buffer. Membranes were collected by centrifugation, dissolved in 1% SDS in boiling water and subjected to gel electrophoresis and autoradiography[2]. The autoradiograph and Coomassie blue-stained gel were both evaluated by densitometry. The relative degree of phosphorylation of spectrin is expressed as the ratio of the areas under the zone profiles, taking that for unextracted membranes to be unity. Some 80–90% of the label was found in the spectrin.

The presence of cAMP in resealed ghosts had no influence on invasion efficiency, which is therefore not related to cAMP-dependent kinase(s). We conclude that a substrate for the cAMP-independent membrane kinase must be phosphorylated if malarial invasion is to occur. This is presumably the basis of the requirement for intracellular ATP[1,2].

Spectrin is by far the most abundantly labelled protein under the conditions of these experiments. On addition of kinase to the extracted cells the enhanced incorporation of label is about 80–90% in spectrin. Given the earlier evidence for direct involvement of the host cell spectrin in the invasion process[1,14], it is a reasonable conjecture (although it cannot be regarded as proven) that spectrin is the invasion-linked substrate for the cAMP-independent kinase.

This investigation received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases. We thank Dr J. C. Pinder for discussion and advice on kinase preparation.

Received 27 February; accepted 30 September 1986.

1. Olson, J. A. & Kilejian, A. *J. Cell Biol.* **95**, 757–762 (1982).
2. Dluzewski, A. R., Rangachari, K., Wilson, R. J. M. & Gratzer, W. B. *Molec. biochem. Parasitol.* **9**, 145–160 (1983).
3. Miller, L. H. *Bull. WHO* **55**, 157–162 (1977).
4. Aikawa, M., Miller, L. H., Johnson, J. & Rabbege, J. *J. Cell Biol.* **77**, 72–82 (1978).
5. Dluzewski, A. R., Rangachari, K., Wilson, R. J. M. & Gratzer, W. B. *Br. J. Haematol.* **55**, 629–637 (1983).
6. Harris, H. W. & Lux, S. E. *J. biol. Chem.* **255**, 11512–11520 (1980).
7. Birchmeier, W. & Singer, S. J. *J. Cell Biol.* **73**, 647–659 (1977).
8. Anderson, J. & Tyler, J. M. *J. biol. Chem.* **255**, 1259–1265 (1980).
9. Patel, V. P. & Fairbanks, G. *J. Cell Biol.* **88**, 430–449 (1981).
10. Loyter, A. A., Ben-Zaquen, R., Marash, R. & Milner, Y. *Biochemistry* **16**, 3903–3909 (1977).
11. Nelson, M. J. & Huestis, W. H. *Biochim. biophys. Acta* **600**, 398–405 (1980).
12. Lalazar, A. & Loyter, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 318–321 (1979).
13. McLaren, D. J., Bannister, L. H., Trigg, P. I. & Butcher, G. A. *Parasitology* **79**, 125–139 (1979).
14. Dluzewski, A. R., Rangachari, K., Wilson, R. J. M. & Gratzer, W. B. *Br. J. Haematol.* **49**, 97–101 (1981).
15. Hosey, M. M. & Tao, M. *Biochim. biophys. Acta* **482**, 348–357 (1977).
16. Tao, M., Conway, R. & Cheta, S. *J. biol. Chem.* **255**, 2563–2568 (1980).
17. Trager, W. & Jensen, J. B. *Science* **193**, 673–675 (1976).
18. Dluzewski, A. R., Ling, I. T., Rangachari, K., Bates, P. & Wilson, R. J. M. *Trans. R. Soc. trop. Med. Hyg.* **78**, 622–637 (1984).
19. Bradford, M. M. *Analyt. Biochem.* **72**, 248–254 (1976).
20. Pinder, J. C., Dhermy, D., Baines, A. J., Lux, S. E. & Gratzer, W. B. *Br. J. Haematol.* **56**, 455–463 (1983).

Plant chitinases are potent inhibitors of fungal growth

Angela Schlumbaum, Felix Mauch,
Urs Vögeli & Thomas Boller

Abteilung Pflanzenphysiologie, Botanisches Institut,
Universität Basel, Hebelstrasse 1, CH-4056 Basel, Switzerland

The antimicrobial arsenal of plants is thought to consist mainly of secondary metabolites, among which the phytoalexins are the best-studied[1–3]. But plants may also possess antimicrobial proteins[4,5]: it has been reported that wheat-germ agglutinin, a chitin-binding lectin from wheat embryos, inhibits growth of the fungus *Trichoderma viride*[4]. This has led to the notion that plant lectins, with their intriguing biochemical similarity to animal antibodies, have an antibody-like antimicrobial function[4,6,7]. We report here that the main proteinaceous inhibitor of fungal growth in bean leaves is chitinase, an enzyme that can be induced by the plant hormone ethylene, or by pathogen attack. Among commercial preparations of purified chitin-binding lectins (from wheat germ, tomato, potato, pokeweed and gorse), only those containing contaminating chitinase activity inhibit fungal growth. Our data indicate that plant chitinases, but not chitin-binding lectins, are important antifungal proteins in plants.

The resistance of plants to potentially pathogenic microorganisms is based on multiple biochemical factors[1]. Particularly important are the phytoalexins, antimicrobial micromolecules that are not present in healthy plants but accumulate in response

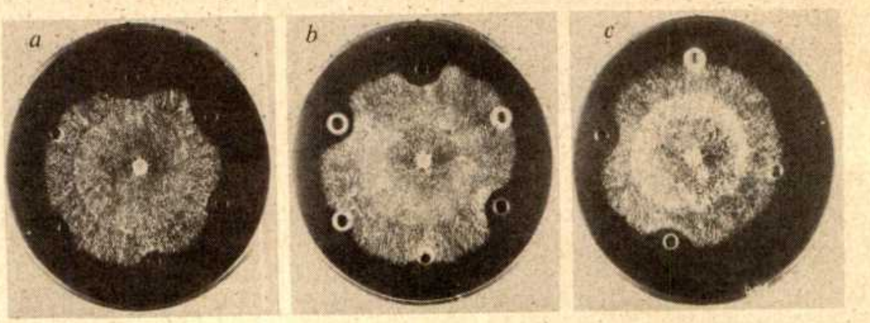


Fig. 1 Growth inhibition of *T. viride* in response to various protein preparations. *a*, Purified bean chitinase[13]. Wells contain 20 μl of the following solutions in μg ml^{-1} (from top, in clockwise direction): 600; 150; 37; 9; 2; 0.6. *b*, Purified chitinase and crude protein extracts from primary bean leaves. Wells contain 20 μl of the following solutions in μg ml^{-1} (from top, in clockwise direction): Purified bean chitinase (130); purified bean chitinase (130) in the presence of antiserum against chitinase; crude protein (12 mg ml^{-1}) from leaves treated with 10 nl ml^{-1} ethylene for 48 h; crude protein (12 mg ml^{-1}) from control leaves; crude protein (12 mg ml^{-1}) from ethylene-treated leaves in the presence of antiserum against chitinase; and crude protein (12 mg ml^{-1}) from ethylene-treated leaves in the presence of preimmune serum. The chitinase activities, determined after the addition of the sera, corresponded to the following concentrations of bean chitinase in μg ml^{-1} (from top, in clockwise direction): 130; 0.1; 130; 1.4; 1.3; 130. *c*, Purified bean chitinase and wheat germ agglutinin (WGA). Wells contain 20 μl of the following solutions (from top, in clockwise direction): commercial wheat germ agglutinin (25 mg ml^{-1}); commercial wheat germ agglutinin (5 mg ml^{-1}); wheat germ agglutinin purified by affinity chromatography on regenerated chitin, 3.5 mg ml^{-1} (with a chitinase activity corresponding to 115 μg ml^{-1} bean chitinase); and bean chitinase, 100 μg ml^{-1}.

Methods. *T. viride* (a gift of Dr Dahmen) was grown on Petri dishes containing 2% malt extract agar at the bottom and 1.5% malt extract agar at the top. One day after inoculation, when the colony diameter was 3-4 cm, wells punched into the 1.5% agar layer 1 cm from the growing front were filled with 20 μl of the various preparations. The plates were photographed 8-12 h later. The crude protein extracts were prepared using 0.1 M potassium phosphate buffer, *p*H 7.0 (4 ml per g fresh weight). After centrifugation (10 min at 10,000 *g*), the supernatant was precipitated overnight with ammonium sulphate at 97% saturation, taken up in half the extraction volume and dialysed twice against 0.01 M potassium phosphate buffer, *p*H 7.0. The extracts were filtered through a 0.22 μm membrane filter (Sartorius, type SM 16534), lyophilized, and taken up in water. Chitinase and protein were measured as previously described[13].

to microbial infection[1-3]. Many of these phytoalexins inhibit growth of fungi at concentrations of 10-50 μg ml^{-1}. In addition, plants may possess proteins with antibiotic activity[4,5]. This was first demonstrated when Mirelman *et al.*[4] found that wheat-germ agglutinin (WGA), a chitin-binding lectin, inhibited growth of the fungus *T. viride* at concentrations of $\geq$600-1200 μg ml^{-1}. Several lectins have since been reported to inhibit fungal growth under similar conditions[8,9]. Plant lectins that bind to fungal cell walls are therefore thought to have an antifungal role[5,6-9].

Another group of plant proteins with potential antifungal functions are the hydrolases that can attack the main components of fungal cell walls, chitin and β-1,3-glucan[1,10]. We have studied chitinase, an enzyme with no known endogenous substrate in higher plants[10,13], which has been purified from various plant tissues, including wheat germ[11], tomato leaves[12] and bean leaves[13]. This enzyme is strongly induced in response to the plant hormone, ethylene[13], and in response to infection by pathogens[12,14,15]. Purified plant chitinase readily liberates chitin oligomers from purified fungal cell walls[12,13] and acts as a lysozyme against bacterial cell walls[10,13]. This is indirect evidence for antimicrobial activity of the enzyme.

To examine antifungal activity directly, we tested the effect of purified bean chitinase on growth of *T. viride*, using the method of Mirelman *et al.*[4]. We observed a marked concentration-dependent inhibition of fungal growth around the wells containing chitinase (Fig. 1*a*). The smallest concentration of chitinase that resulted in a distinct inhibition zone was 2 μg ml^{-1}. This means that chitinase is 300 times as potent as the WGA preparation of Mirelman *et al.*[4].

Boiled chitinase had no effect on fungal growth. Most plant chitinases are basic proteins[10], but another basic protein, cytochrome *c*, did not inhibit fungal growth in the bioassay at concentrations up to 10 mg ml^{-1} (data not shown). When antibodies were raised against purified bean chitinase in a rabbit, the antiserum was found to block chitinase activity. It also blocked the inhibitory effect of chitinase on fungal growth (Fig. 1*b*) whereas preimmune serum of the same rabbit had no such effect.

We tested crude protein extracts from bean leaves in the fungal growth assay, using a protein concentration (12 mg ml^{-1}) equivalent to that in leaves (Fig. 1*b*). Protein extracts from ethylene-treated leaves produced much larger inhibition zones than extracts from control leaves. This correlates well with strong induction of chitinase by ethylene[13]. The crude extracts from ethylene-treated leaves had a chitinase activity equivalent to 130 μg ml^{-1} pure chitinase. The same amount of pure chitinase produced an inhibition zone of a similar size. The inhibitory effect of the crude extract was almost completely overcome by the addition of antiserum against chitinase, whereas preimmune serum did not affect it (Fig. 1*b*). These data indicate that chitinase is the main inhibitor of fungal growth in crude protein preparations from bean leaves.

We have also examined bean leaves from plants infected with bean rust, *Uromyces phaseoli*, using the bean cultivar Warox, which is hypersensitively resistant to the pathogen used. We found that chitinase was about tenfold induced in response to infection. Crude extracts from infected leaves produced larger inhibition zones than did extracts from uninfected control leaves; again, adding the antiserum against bean chitinase prevented the formation of an inhibition zone (Table 1).

Table 1 Inhibtition of fungal growth by crude extracts from rust-infected and uninfected bean leaves

Test Preparation	Protein mg ml^{-1}	Chitinase μg ml^{-1}	Inhibition of Fungal Growth
Extract from	7.0	3.3	+
uninfected leaves	3.5	1.6	−
Extract from	6.6	30	+++
infected leaves	3.3	15	+++
	1.7	7.5	++
Extract from infected leaves and antiserum	6.6	<0.05	−

Bean seedlings (c. Warox, supplied by van Waveren Pflanzenzucht, Rosdorf, FRG) were sprayed 10-14 days after germination with 1% Tween-80 (controls) or with 1% Tween-80 containing 6×10^4 spores of bean rust per ml (*U. phaseoli*, supplied by Dr U. Gisi). The plants were kept for 48 h at 100% humidity and further incubated in a greenhouse for 12 days. Protein extracts from primary leaves were prepared as described in the legend to Fig. 1. Chitinase activity[13] is expressed in equivalents of pure bean chitinase[15]. The effect on growth of *T. viride* (tested with 20 μl of each preparation as in Fig. 1) is expressed in qualitative terms: +++ stands for strong inhibition, + for just detectable inhibition.

Table 2 Effect of chitin-binding plant lectins on fungal growth

Lectin preparation	Protein mg ml^{-1}	Titre μg ml^{-1}	Chitinase μg ml^{-1}	Inhibition of fungal growth
Tomato lectin*	2	<0.5	13	+++
Potato lectin*	2	0.9	1.7	+
Pokeweed lectin†	5	30	0.7	++
Pokeweed lectin*	3	0.2	<0.05	–
Gorse lectin UEA II*	2	60	0.06	–

The titre for the agglutination of trypsinized red blood cells and the chitinase activity[13], expressed in equivalents of pure bean chitinase[15], were determined for each lectin preparation. The effect on growth of *T. viride* (tested with 20 μl of each preparation as in Fig. 1) is expressed in qualitative terms as in Table 1.

* Obtained from Sigma.

† Obtained from Serva.

To reexamine the antifungal effect of WGA[4], we verified that purified WGA (Serva) agglutinated trypsinized red blood cells (titre: 2.4 μg ml^{-1}) and found that it contained no measurable chitinase activity. This WGA preparation did not inhibit growth of *T. viride* even at concentrations of 25 mg ml^{-1} (Fig. 1*c*). Preparations of WGA from Sigma, Calbiochem and Pharmacia were similarly free of contaminating chitinase and did not inhibit fungal growth.

Mirelman *et al.*[4] purified WGA by affinity chromatography on sepharose-2-acetamido-N-(5-aminocaproyl)-2-deoxy-β-D-glucopyranosylamine. However, affinity chromatography of WGA may result in copurification of WGA and chitinase[11]. WGA prepared by affinity chromatography on regenerated chitin[11,16] (a readily available affinity matrix) agglutinated trypsinized red blood cells with a titre of 3.3 μg ml^{-1} and contained contaminating chitinase activity equivalent to ~40 μg bean chitinase per mg protein. (Wheat-germ chitinase had previously been purified from a similar affinity-purified WGA preparation[11].) This chitinase-contaminated WGA preparation strongly inhibited growth of *T. viride* at a concentration of 3.5 mg ml^{-1} (Fig. 1*c*). The antiserum against bean chitinase did not inhibit the chitinase activity in WGA and did not reduce the size of the inhibition zone caused by the contaminated WGA. When pure bean chitinase was added to chitinase-free commercial WGA, the combination inhibited fungal growth in the same way as affinity-purified WGA (data not shown). Thus, the inhibitory effect of WGA observed by Mirelman *et al.*[4] may have been caused by contamination with chitinase.

We have also tested the effect of several other commerical lectins with a specificity for N,N'-diacetylchitobiose or N,N',N''-triacetylchitotriose (Table 2). The tomato and the potato lectin preparations contained contaminating chitinase activity. Both inhibited fungal growth. The chitinases present in these preparations were found to be serologically related to bean chitinase as the antiserum against bean chitinase totally inhibited the chitinase activity of both. The antiserum also suppressed the growth-inhibiting activity of the lectin preparations completely (data not shown). We obtained two different preparations of pokeweed (*Phytolacca americana*) lectin. One preparation (Serva), had lower lectin activity than the other. It was contaminated with chitinase and inhibited fungal growth (Table 2). The other preparation (Sigma), although more active as a lectin, did not contain chitinase and did not inhibit fungal growth. A gorse agglutinin, UEA II, contained almost no chitinase activity and did not inhibit fungal growth. Collectively, these data show that preprations of chitin-binding lectins inhibit growth of *T. viride* only when they are contaminated by chitinase. Of further commercial lectin preparations with other specificities examined (concanavalin A, bean phytohemagglutinin A, peanut agglutinin, osage orange lectin), none was contaminated by chitinase, and none inhibited fungal growth.

We have demonstrated directly that plant chitinases are highly potent growth inhibitors of *T. viride*. It is likely that the effect is attributable to enzymatic attack on newly formed chitin in the growing hyphal tips[11] disturbing the balance between synthesis and hydrolysis required for tip growth[17]. It is intriguing that chitinase and β-1, 3-glucanase are coordinately induced by ethylene or by pathogen attack in plants[10,14]. Chitinase alone, at concentrations of 2 μg ml^{-1}, is sufficient to stop growth of *Trichoderma viride*, a non-pathogenic test fungus. Inhibition of several other fungi, however, requires the presence of a combination of chitinase and β-1, 3-glucanase at similar concentrations (F.M., B. Mauch and T.B., in preparation). On a concentration basis, antimicrobial plant hydrolases, either alone or in combination, are inhibitors of fungal growth at least as potent as the phytoalexins. They may be similarly important in the defences of plants against pathogens.

This work was supported by grant 3.400.0.83 of the Swiss National Science Foundation.

Received 22 July 1986; accepted 30 September 1986.

1. Bell, A. A. *A. Rev. Pl. Physiol.* **32,** 22-81 (1981).
2. Bailey, J. A. & Mansfield J. W. (eds) *Phytoalexins* (Halsted, Wiley, New York, 1982).
3. Darvill, A. G. & Albersheim, P. *A. Rev. Pl. Physiol.* **35,** 243-275 (1984).
4. Mirelman, D., Galun, E., Sharon, N. & Lotan, R. *Nature* **256,** 414-416 (1975).
5. Roberts, W. K. & Selitrennikoff, C. P. *Biochim. biophys. Acta* **880,** 161-170 (1986).
6. Sequeira, L. *A. Rev. Phytopathol.* **16,** 453-481 (1978).
7. Etzler, M. E. *A. Rev. Pl. Physiol.* **36,** 209-234 (1985).
8. Barkai-Golan, R., Mirelman, D. & Sharon, N. *Archs. Microbiol.* **116,** 119-124 (1978).
9. Brambl, R. & Gade, W. *Physiologia Pl.* **64,** 402-408 (1985).
10. Boller, T. in *Cellular and Molecular Biology of Plant Stress* (eds J. L. Key & T. Kosuge), 247-262 (Liss, New York, 1985).
11. Molano, J., Polacheck, I., Duran, A. & Cabib. E. *J. biol. Chem.* **254,** 4901-4907 (1979).
12. Pegg, G. F. & Young, D. H. *Physiol. Pl. Pathol.* **21,** 389-409 (1982).
13. Boller, T., Gehri, A., Mauch, F. & Vögeli, U. *Planta* **157,** 22-31 (1983).
14. Mauch, F., Hadwiger, L. A. & Boller, T. *Pl. Physiol.* **76,** 607-611 (1984).
15. Métraux, J. P. & Boller, T. *Physiol. molec. Pl. Pathol.* **28,** 161-169 (1986).
16. Bloch, R. & Burger, M. M. *Biochem. biophys. Res. Commun.* **58,** 13-19 (1974).
17. Farkas, V. *Microbiol. Rev.* **43,** 117-144 (1979).

Sulphonolipids are molecular determinants of gliding motility

Darren R. Abbanat, Edward R. Leadbetter, Walter Godchaux III & Alan Escher

Department of Molecular and Cell Biology, U-131,
University of Connecticut, Storrs, Connecticut 06268, USA

A variety of bacteria[1] that possess no obvious locomotor organelles are, nonetheless, able to translocate on solid surfaces (but not through liquids). In no case has the mechanism of this 'gliding' motility been elucidated[2]. Our recent discovery that unusual sulphonolipids[3,4] (Fig. 1) are major cell-envelope components peculiar to the *Cytophaga-Flexibacter* group of gliding bacteria[5] led us to examine whether these lipids are important in this motility. Mutants deficient in both gliding and sulphonolipid synthesis were isolated from mutagenized cultures of *Cytophaga johnsonae* (our laboratory strain, originally ATCC 17061). In some of these mutants, restoration of sulphonolipid content through provision of a specific biosynthetic precursor resulted in recovery of the ability to glide. The sulphonolipids are the first molecules shown to be specifically required for gliding motility (surface translocation).

Cultures ($A_{650} = 0.8$) mutagenized with *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (0.3 mg ml^{-1}) for 60 min were diluted and plated on LTY agar (Fig. 2 legend); 48 spreading-deficient clones were selected on the basis of colony morphology. The colonies formed by gliding cells expand rapidly, are thin and have finger-like projections at their edges, whereas those of nonmotile cells are compact, rounded and have sharply-demarcated edges (Fig. 2). In minimal liquid medium (fructose plus inorganic salts, with sulphate as the sole sulphur source), all the mutant strains considered below grew with the same doubling time (±5%) as

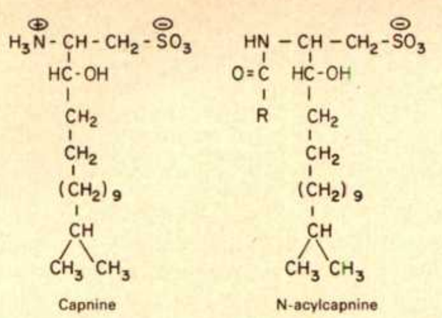


Fig. 1 Structures of the sulphonolipids. R–CO– represents one of a variety of long-chain fatty acids[4]. *N*-acylcapnine is the predominant species found in the parent strain of *C. johnsonae*; capnine is apparently a precursor of the acylated form.

the parent strain, so their abnormal colony morphology was not the result of growth defects. These spreading-deficient strains were examined for sulphonolipid content. Cells were grown in minimal medium containing $Na_2^{35}SO_4$ and lipid radioactivity was measured after separation by thin-layer chromatography (methodological details are given in ref. 5). Four strains were found to be deficient in some aspect of their sulphonolipid content.

Mutant 21A2I contained less than 2% of the amount of sulphonolipid found in the parent strain (Table 1) and was a virtual nonglider, as judged not only by the failure of its colonies to spread (Fig. 2, top row) but also by direct observation using the phase-contrast microscope. Gliding cells form 'rafts'—groups of cells gliding together—that can be observed when they detach from the colony edge. In the parent strain these rafts moved at an average rate of about 25 μm min^{-1} (Table 1). In contrast, the edges of colonies of mutant 21A2I were smooth and no gliding movement of cells could be observed.

We have shown[6] that L-cysteate is the immediate precursor of carbons 1 and 2, the sulphonate group and (probably) the amino group of capnine (Fig. 1). Cysteate is thought to condense with an activated fatty acid to form, ultimately, capnine[6], in a manner analogous to the biosynthesis of sphingosine from serine and a fatty acyl coenzyme A[7]; capnine is then *N*-acylated. Provision of L-cysteate in the growth medium restored the sulphonolipid content of mutant 21A2I to normal (Table 1); as is the case in the parent strain[5], virtually all the sulphonolipid thus formed was in the form of *N*-acylcapnine. Provision of L-cysteate also restored gliding motility; colony spreading became normal (Fig. 2, bottom row) and rafts were formed and moved at about the same speed as those of the parent strain (Table 1). Evidently mutant 21A2I is unable to form cysteate, an obligatory precursor of capnine, from the standard sulphur sources provided in these experiments but when given cysteate it can synthesize sulphonolipid and therefore can glide. A second mutant, number 42, is apparently a 'leaky' version of 21A2I. It contains detectable sulphonolipid but only 13% of normal levels; its colony-spreading and raft-movement rates are also low (70 and 30% of the parent-strain rates, respectively) and L-cysteate restores gliding ability to normal.

The effects of cysteate are stereospecific; D-cysteate, which does not compete with L-cysteate for incorporation into capnine (unpublished observations), did not restore gliding motility (Fig. 2, middle row). Hence, the effect of L-cysteate on gliding must be a biological one and not some gross physical effect of a sulphonate in the medium. A number of other compounds tested, including L-homocysteate, taurine, L- and D-cystine, L-cysteine sulphinic acid, L-methionine, L-djenkolate and 2-mercaptoethanesulphonic acid (coenzyme M), proved equally ineffective in restoring gliding. We also note that *C. johnsonae* cannot use cysteate as the sole sulphur source (unpublished observation) and that sulphur and carbon from exogenous L-cysteate are incorporated into sulphonolipids but not into protein[6]; indeed, the compound appears not to be metabolized except by incorporation into capnine. Thus it is not likely that cysteate restores gliding via some metabolic transformation other than conversion to sulphonolipid. These observations support the conclusion that sulphonolipids are necessary for gliding and militate strongly against other interpretations, as for example that 21A2I is a double mutant, or that it is deficient in some cellular product that is independently necessary both for sulphonolipid synthesis and for gliding. Such interpretations require two distinct roles for cysteate in restoring both sulphonolipid content and gliding motility.

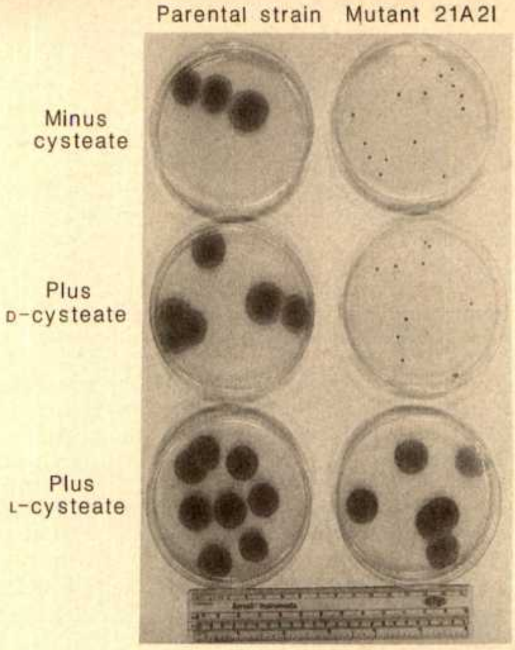


Fig. 2 Colony spreading in *C. johnsonae*. Cells of the parent strain or mutant 21A2I (as indicated) were cloned on LTY medium (0.05% yeast extract, 0.05% Tryptone, solidified with 1.8% agar) containing (where indicated) D- or L-cysteate (25 μM). After growing for 84 h at 30 °C, the colonies were stained with tetrazolium blue to enhance contrast for photography. A six-inch ruler is shown for scale.

Two additional gliding-deficient, sulphonolipid-deficient mutants were studied. Number 36 is phenotypically identical to 21A2I in mutant colony spreading and raft movement. It has normal total lipid sulphur content, over 99% of this in the form of free (non-acylated) capnine and there is no detectable *N*-acylcapnine (whereas in the parent strain, the converse is true). As expected, provision of L-cysteate does not restore gliding.

Table 1 Properties of mutant 21A2I

Strain	Addition to growth medium	Raft movement (μm min^{-1})	Sulphonolipid content (μmol per g cells, wet wt)
Parent	None	25±6	6.2±1.2
Parent	L-cysteate	24±8	10.0±2.6
21A2I	None	None	<0.1
21A2I	L-cysteate	35±8	9.4±2.6

Where indicated, L-cysteate was present at 25 μM. Raft movement was measured directly, on colonies grown as described in Fig. 2 legend, using a phase-contrast microscope fitted with a television monitor. Values are the mean (±standard deviation) of at least 12 independent determinations; 21A2I showed no raft formation or gliding movement in unsupplemented medium. When cysteate was absent, sulphonolipid content was measured, as described in the text, using $Na_2^{35}SO_4$ as the sole sulphur source. Cysteate, when added, was labelled with ^{35}S. Values are the mean (±s.d.) of at least six independent determinations.

Evidently, mutant 36 can synthesize cysteate and capnine but is defective in the *N*-acylation of capnine. This suggests that ability to form the *N*-acylated lipid is the specific requirement for gliding, but this conclusion must be regarded as tentative in the absence of supporting evidence. Another mutant, number 15, glides slowly (raft-movement and colony-spreading rates are 28 and 40% of the parent-strain values, respectively). Its lipid sulphur content is 40% of that found in the parent strain; it has traces of both capnine and *N*-acylcapnine, but most of the sulphur is found in abnormal sulphonolipids that do not co-chromatograph with either. Their nature is under investigation.

As has been reported[8] for many other non-gliding, non-spreading mutants of this bacterium, cells of the four mutants reported here display several behaviours characteristic of wild-type cells when examined microscopically in wet mounts of suspensions on glass slides. These include cell movements described[9] as spinning and pivoting, latex microsphere (260 nm diameter) movement over cell surfaces, and brief translocational movements of the cells (which, in the case of our mutants and parent strain, rarely exceed a single cell length on glass surfaces). We therefore conclude that sulphonolipids, although required for gliding motility and colony spreading on agar, are not required for the movements of *C. johnsonae* seen in wet mounts.

It seems reasonable to suppose that the cell and cell-surface movements observed in the wet mounts are caused by some subset of the machinery responsible for biologically significant gliding motility (cell translocation over long distances on solid surfaces). At present, however, the evidence[8-11] for this is indirect and the machinery remains undescribed. Our demonstration of the importance of the novel sulphonolipids in gliding motility provides new opportunities for examining the molecular basis of that process in these bacteria and for clarifying the relationship between cell movements in wet mounts and the mass translocation of cells over solid surfaces.

This research was supported by NSF grant PCM8214050.

Received 15 April; accepted 15 September 1986.

1. Reichenbach, H. *A. Rev. Microbiol.* **35**, 339-364 (1981).
2. Burchard, R. P. *A. Rev. Microbiol.* **35**, 497-529 (1981).
3. Godchaux, W. & Leadbetter, E. R. *J. Bact.* **144**, 592-602 (1980).
4. Godchaux, W. & Leadbetter, E. R. *J. biol. Chem.* **259**, 2982-2990 (1984).
5. Godchaux, W. & Leadbetter, E. R. *J. Bact.* **153**, 1238-1246 (1983).
6. Abbanat, D. R., Godchaux, W., Polychroniou, G. & Leadbetter, E. R. *Biochem. biophys. Res. Commun.* **130**, 873-878 (1985).
7. Brady, R. O. & Koval, G. O. *J. biol. Chem.* **233**, 26-31 (1958).
8. Chang, L.-Y. E., Pate, J. L. & Betzig, R. O. *J. Bact.* **159**, 26-35 (1984).
9. Pate, J. L. & Chang, L.-Y. E. *Curr. Microbiol.* **2**, 59-64 (1979).
10. Wolkin, R. H. & Pate, J. L. *J. gen. Microbiol.* **131**, 737-750 (1985).
11. Lapidus, I. R. & Berg, H. C. *J. Bact.* **151**, 384-398 (1982).

Ion channels in human neutrophils activated by a rise in free cytosolic calcium concentration

V. von Tscharner, B. Prod'hom*, M. Baggiolini & H. Reuter*†

Theodor Kocher Institute and * Department of Pharmacology,
University of Berne, 3010 Berne, Switzerland

A rapid, transient rise in the free cytosolic Ca^{2+} concentration ($[Ca^{2+}]_i$) is one of the earliest events in neutrophil activation and is assumed to be involved in many of the subsequent cellular reactions[1-3]. Both Ca^{2+} release from intracellular stores and Ca^{2+} influx from the extracellular space contribute to the rise in $[Ca^{2+}]_i$[4]. In an attempt to assess the relative importance of these pools and the sequences leading to the rise in $[Ca^{2+}]_i$, we have studied the time course of changes in $[Ca^{2+}]_i$ after stimulation with *N*-formyl-methionyl-leucyl-phenylalanine (fMLP) or platelet-activating factor (PAF) using the Ca^{2+} indicators quin-2 and fura-2. We observed a time lag of 1–3 s between stimulation and rise in $[Ca^{2+}]_i$. This lag depends on the agonist concentration but is independent of extracellular Ca^{2+}. Thus Ca^{2+} release from intracellular stores is rate limiting for the rise in $[Ca^{2+}]_i$. After this, cation channels in the plasma membrane (measured with the patch clamp method[5]) are opened. These non-selective channels, which also pass Ca^{2+}, are activated by the initial rise in $[Ca^{2+}]_i$, but by neither fMLP nor inositol 1,4,5-trisphosphate (IP_3) directly.

The time courses of change in $[Ca^{2+}]_i$ were measured in purified human neutrophils loaded with the indicators quin-2 (ref. 6) or fura-2 (ref. 7). Shortly before stimulation the cells were suspended in an isotonic saline solution containing either 1 mM $CaCl_2$ or 1 mM EGTA with no $CaCl_2$. Figure 1 shows the increase in $[Ca^{2+}]_i$ measured with quin-2 after stimulation of the neutrophils with the calcium ionophore ionomycin or the chemotactic peptide fMLP. A rise in $[Ca^{2+}]_i$ was observed in both cases in the presence and absence of extracellular Ca^{2+}. The time courses, however, were different: the fluorescence changes induced by ionomycin started in less than 0.2 s, whereas those induced by fMLP were delayed up to 3 s ($\sim$1 s with fMLP 10^{-7} M; Fig. 1). The rate of fluorescence increase following either stimulus was about twice as fast when extracellular Ca^{2+} was present, because Ca^{2+} influx from the extracellular space is added to Ca^{2+} release from intracellular stores[8-10]. The experiments with ionomycin show that: (1) $[Ca^{2+}]_i$ changes mediated by the ionophore occur without a measurable time lag and (2) the time resolution of our testing system (depending on mixing, Ca^{2+} buffering by the indicator and the response time of the apparatus) was well above the range of the lag periods after stimulation with fMLP. Thus the delay in the response to fMLP reflects the time required for stimulus transduction.

We performed further experiments using fura-2 as the indicator and either fMLP or PAF as stimuli. Fura-2 can be used at lower concentrations than quin-2, resulting in less buffering of $[Ca^{2+}]_i$ because of its higher fluorescence quantum yield[7]. The fura-2 fluorescence changes were comparable to those seen with quin-2. A lag period of similar duration was observed between stimulation with either fMLP or PAF and the increase in $[Ca^{2+}]_i$. The lag time shortens and the rate of fluorescence change increases markedly with increasing concentrations of fMLP (Fig. 2). Similar families of curves were recorded upon stimulation with PAF (not shown). With either stimulus the duration of the lag was virtually identical with or without extracellular Ca^{2+} (6 experiments).

Thus the rise in $[Ca^{2+}]_i$ is initiated by a rate-limiting process that does not depend on extracellular Ca^{2+}. But it is clear that Ca^{2+} influx does occur in fMLP- or PAF-stimulated cells (Fig. 1)[8-10]. What controls the Ca^{2+} permeability of the plasma membrane? We have performed patch clamp experiments[5] and have measured ion channel activity in the neutrophil plasma membrane directly. Channels with similar properties were recorded in both cell-attached and isolated, inside-out membrane patches. In resting cells the occurrence of channel openings was extremely rare and the opening probability was not increased in the presence of fMLP (10^{-8} to 10^{-6} M) in the pipette solution. This strongly suggests that the channels are not directly coupled to the agonist receptor. If fMLP was added to the bathing medium with Ca^{2+} concentrations of 1-2 mM, about half (13/24) of the cells responded with increased channel openings. If cells were loaded with fura-2 in the absence of external Ca^{2+} (ref. 11) $[Ca^{2+}]_i$ was reduced to about 10^{-9} M and fMLP did not induce channel openings. Extensive single channel activity was observed in inside-out membrane patches, provided the Ca^{2+} concentration in the bathing solution (which is in contact with

---

† To whom correspondence should be addressed.

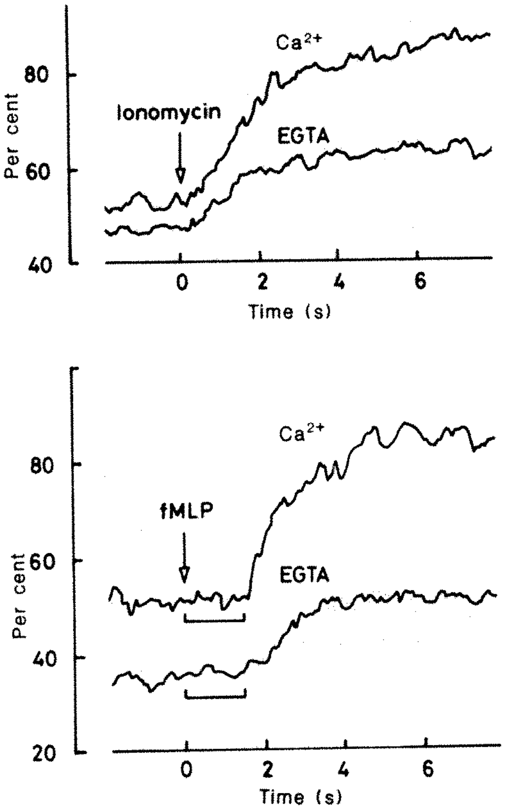


Fig. 1 Quin-2 fluorescence curves in human neutrophils following addition (arrow) of ionomycin (2×10^{-6} M) or fMLP (10^{-7} M). The upper traces (Ca^{2+}) were obtained in the presence of 1 mM $CaCl_2$ in the medium, and the lower traces (EGTA) in a calcium-free medium with 1 mM EGTA added 5 min before stimulation. The lag times between fMLP addition and onset of quin-2 fluorescence increase are indicated by the bars underneath the tracings. **Methods.** Neutrophils were obtained from buffy coats of donor blood stored for up to 24 h at 4–10 °C (Swiss Red Cross Laboratory, Berne, Switzerland) and purified by centrifugation through a discontinuous Lympho-paque gradient[21]. The cells collected at the boundary between the two density layers (95–98% neutrophils) were washed three times by centrifugation (160*g* for 5 min at room temperature) in a saline medium containing 136 mM NaCl, 4.8 mM KCl, 1.2 mM KH_2PO_4, 5 mM glucose, buffered at *p*H 7.4 with 20 mM *N*-2-hydroxyethylpiperazine-*N*-2-ethanesulphonate (HEPES/NaOH). The final suspension was adjusted to 4×10^6 cells per ml and kept at room temperature until use. Loading with the fluorescent indicator: Neutrophils (4×10^6 cells per ml) were incubated for 30 min at 37 °C in the presence of quin-2 acetoxymethyl ester (0.4 nmol per 10^6 cells) dispensed from a 2 mM stock solution in dimethyl sulphoxide. The process was followed fluorimetrically and was considered to be complete when a maximum was observed in the 510/465 (nm/nm) fluorescence intensity ratio (510 nm is just above the fluorescence peak; 465 nm is an arbitrary reference point that changes very little as loading proceeds). A filter fluorimeter was specially constructed for high time resolution measurements. The excitation wavelength was selected from a 500 W Osram Hg-lamp with a band filter of 334 ± 10 nm. A 470 nm cut-off filter was inserted on the emission side allowing the measurement of the whole emission band with a photomultiplier (EMI 9798QB) connected to an amplifier with a response time of 30 ms. Broad band emission measurement and the indicated amplifier's response time resulted in a reasonably high signal-to-noise ratio. Loaded neutrophils were stimulated at 37 °C in disposable polystyrol cuvettes (8 mm inner diameter) under magnetic stirring. The agonist was injected through a light-tight opening in the lid of the cuvette holder with a Hamilton syringe. Injection was timed by a triggering signal to the recorder. Intracellular calibration of the quin-2 signal was used to determine the quin-2 saturation fraction *q*. Ionomycin (2×10^{-6} M) was added to the cells in the presence of 1 mM extracellular Ca^{2+} to bring about complete saturation of intracellular quin-2 and the fluorescence was taken as F_{max}. All quin-2 fluorescence was then quenched by addition of 10 mM $MnCl_2$ and the background fluorescence was taken as F_0. In agreement with earlier work[22] the fraction of total quin-2 present as calcium complex *q* is then calculated according to the relations:

$$q=\frac{(F-F_{min})}{(F_{max}-F_{min})}$$

where

$$F_{min}=F_{max}-\tfrac{5}{6}(F_{max}-F_0)$$

the inner surface of the patch) was higher than 10^{-7} M. We have also measured the activity of these channels in saponin treated cells. We exposed whole cells to low (0.005%) saponin concentrations for ~10 min after establishing a gigaohm seal between pipette tip and membrane[12]. Saponin makes membranes permeable to small molecules. If Ca^{2+} was present in the external medium (10^{-5}–10^{-3} M) we observed activation of ion channels identical to those seen in inside-out patches, but if the Ca^{2+} concentration was lowered to 10^{-8} M the activity disappeared. Two types of single channel currents could be identified (Fig. 3). Although we have no direct evidence that the two current levels in Fig. 3 result from openings of two distinct populations of ion channels, in experiments with excised inside-out patches the larger currents disappeared more rapidly (within a few minutes) than the smaller ones, transitions between current levels were not strictly coupled and in many patches only one type of current could be seen. Therefore, we suggest that the currents flow through two separate types of ion channels.

With identical solutions (130 mM KCl or K-aspartate) on both sides of an inside-out membrane patch the current-voltage relationships were roughly linear with slope conductances of 18–25 pS and 4–6 pS at about 24 °C, respectively. Both channels discriminate only poorly between Na^+, K^+ and Ca^{2+}. This was assessed by measuring the reversal potential of single channel currents with two different ion species on either side of excised inside-out patches of membrane (Fig. 4*a*,*b*). With 90 mM $CaCl_2$ solution in the pipette and 130 mM KCl or K-aspartate solution in the bath, the reversal for both channel currents was near 0 mV (Fig. 4*d*), indicating that both channels were permeable to Ca^{2+} and K^+. The same was true when a 130 mM NaCl solution replaced the K-aspartate solution in the bath or the $CaCl_2$ solution in the pipette. Outward currents through the channels at positive membrane potentials were usually larger than inward currents (Fig. 4*d*). Channel openings were seen more often at positive than at negative membrane potentials, implying some voltage dependence of channel gating. However, single channel currents could only be measured if the Ca^{2+} concentration in the bath was raised to above 10^{-7} M (Figs 3*c* and 4*c*). Therefore, a rise in $[Ca^{2+}]_i$ is an essential step in the gating of both channels. Open times ranged from a few milliseconds up to several seconds. The smaller conductance channel was more likely to stay open for many seconds. IP_3, which has been shown to release Ca^{2+} from interacellular stores in neutrophils[13], does not by itself induce ion channel activity when added to the internal surface of the membrane (Fig. 3*b*).

Our findings strongly suggest that in neutrophils the first step leading to a change in $[Ca^{2+}]_i$ is the release of Ca^{2+} from intracellular stores. This Ca^{2+} release seems to be induced by IP_3 (ref. 13) formed after agonist-receptor occupancy[14]. The increase in $[Ca^{2+}]_i$ due to Ca^{2+} release causes transient openings of Ca^{2+}-gated ion channels in the membrane through which more Ca^{2+} ions flow into the cytoplasm. Over a time period of 1-2 min the fluorescence signals decrease again (not shown) indicating a reduction in the free cytoplasmic Ca^{2+} concentra-

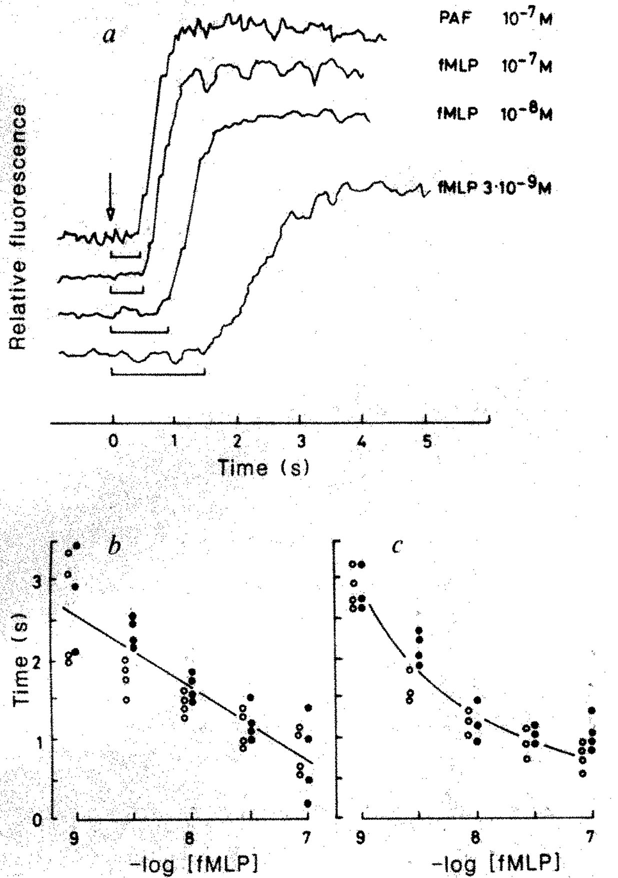


Fig. 2 *a*, Fura-2 fluorescence increase in human neutrophils stimulated with PAF or fMLP. The lag times between agonist addition (arrow) and the onset of fluorescence increase are indicated by the bars under the single tracings. Average $[Ca^{2+}]_i$ prior to stimulation with agonists was 1.3×10^{-7} M. *b*, Plot of the lag time between fMLP addition and onset of fluorescence increase against fMLP concentration. Single values obtained with the same cell preparation in the presence (●) and absence (○) of extracellular calcium are shown with the calculated regression line. No statistical difference was observed between the regression lines through the values obtained with or without extracellular calcium. *c*, Time for an increase in fluorescence to 90% of maximum against concentration of fMLP. Data from the same series of experiments as *b*.

Methods. The neutrophils were prepared as described (Fig. 1) and loaded with fura-2 (0.1 nmol per 10^6 cells) using the same protocol as for quin-2. Fluorescence measurements were made with the instrument described in Fig. 1 legend. Excitation was at 334 ± 10 nm and emission was recorded through a band-pass filter of 500 ± 10 nm. Since lag time measurements do not require it, absolute calibration of the fura-2 signal was omitted to allow frequent repeats of the assay with single batches of cells. A fluorescence decrease due to degranulation as seen in mast cells[23] was not observed.

tion. In intact cells we found spontaneous inactivation of the cation channels which may contribute to this fall in $[Ca^{2+}]_i$. Since the opening of the poorly selective cation channels is not rate limiting for the initial rise in $[Ca^{2+}]_i$, the observed lag period is independent of the external Ca^{2+} concentration. Influx of Na^+ in addition to Ca^{2+} ions through these channels accounts for the depolarization associated with neutrophil stimulation[15]. We found no evidence for voltage-dependent Ca^{2+} channels as in neoplastic B lymphocytes[16]. The channels described in this paper also differ from other Ca^{2+}-activated non-selective cation channels[17,18] in being highly permeable to Ca^{2+} ions. Although the exact role of Ca^{2+} in other neutrophil reactions (for example, superoxide anion production, degranulation, aggregation) is still

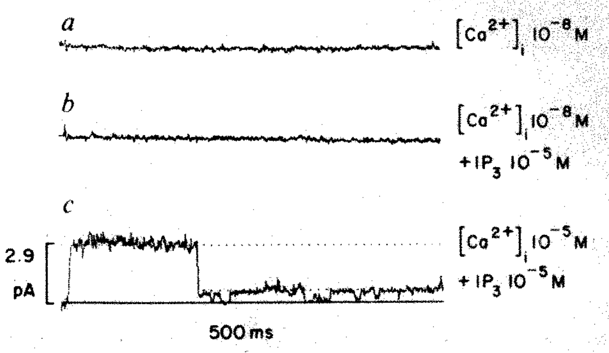


Fig. 3 Ion channel activation by $[Ca^{2+}]_i$ in an inside-out patch of neutrophil membrane. *a*, 10^{-8} M $[Ca^{2+}]_i$; *b*, 10^{-8} M $[Ca^{2+}]_i + 10^{-5}$ M IP_3; *c*, 10^{-5} M $[Ca^{2+}]_i + IP_3$. No channel openings were seen in *a* or *b*, but two outward current levels resulting from openings of ion channels can be seen in *c*. Presence of IP_3 was not required for activation of the channels. The membrane potential of the patch was held at −100 mV and currents were recorded during voltage steps of 500 ms duration to +100 mV.

Methods. Single channel currents were recorded in an excised, inside-out patch of membranes from human neutrophils using the patch clamp method[5]. The bath solution (23-25 °C) contained 130 mM K-aspartate, 20 mM HEPES buffer (*p*H 7.4) and 10^{-8} M $[Ca^{2+}]_i$ (0.32 mM $CaCl_2$ + 1 mM EGTA; apparent stability constant $10^{7.5}$ M^{-1}), or 10^{-5} M (unbuffered) free $[Ca^{2+}]_i$; the solution in the recording pipette contained 90 mM $CaCl_2$ and 10 mM HEPES buffer (*p*H 7.4). Digitized records (1 ms per point) were analysed by a computer (PDP 11/04). Capacity transients and leakage currents during voltage steps have been subtracted. Heavy lines in the records indicate the closed state of the channel while dotted lines show the average current levels when the channels are open.

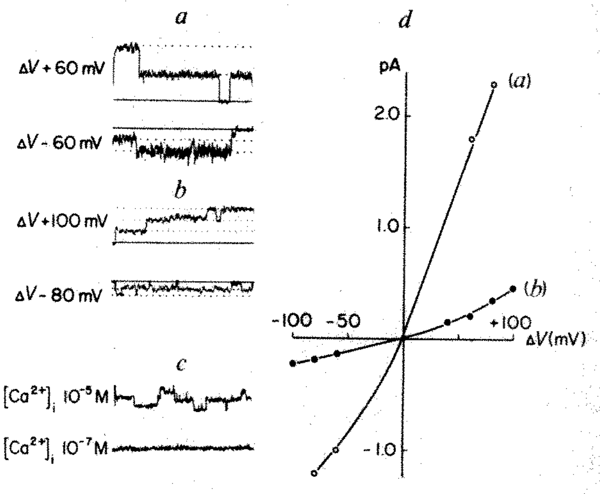


Fig. 4 Examples of current traces (*a*–*c*) and corresponding current–voltage relationships (*d*) of two types of ion channels in excised inside-out membrane patches from human neutrophils. Bath and pipette solutions were the same as in Fig. 3. Outward currents through open channels (upward deflections from heavy base line) were carried by K^+ ions, while inward currents (downward deflections) were carried by Ca^{2+} ions (*a*, *b*). Channel openings were observed only when the $[Ca^{2+}]$ at the inner surface of the membrane patch was raised above 10^{-7} M (*c*). Time scales are 0.5 s in *a* and *b* and 10 s in *c*. Current amplitudes during single channel openings at different membrane potentials have been plotted in *d*. ○, Current-voltage relationship from single-channel currents in *a*; ●, current-voltage relationship from single-channel currents in *b*. Note that outward currents at positive pulse potentials (Δ*V*) are larger than inward currents at negative potentials. The reversal potential is near 0 mV, indicating similar permeabilities for K^+ and Ca^{2+} through the channel.

unclear[19], the rise in $[Ca^{2+}]_i$ appears to be an important intermediary step in neutrophil activation[15,20].

We thank the Swiss National Science Foundation (grants 3.353-0.82 to M.B. and 3.059-0.84 to H.R.) for generous support and Markus Gerber for technical assistance.

Received 21 April; accepted 24 September 1986.

1. Pozzan, T., Lew, D. P., Wollheim, C. B. & Tsien, R. Y. *Science* **221**, 1413-1415 (1983).
2. White, J. R., Naccache, P. H., Molski, T. F. P., Borgeat, P. & Sha'afi, R. I. *Biochem. biophys. Res. Commun.* **113**, 44-50 (1983).
3. Gennaro, R., Pozzan, T. & Romeo, D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1416-1420 (1984).
4. Korchak, H. M. *et al. J. biol. Chem.* **259**, 4076-4082 (1984).
5. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85-100 (1981).
6. Tsien, R. Y. *Biochemistry* **19**, 2396-2404 (1980).
7. Grynkiewicz, G., Poenie, M. & Tsien, R. Y. *J. biol. Chem.* **260**, 3440-3450 (1985).
8. Sklar, L. A. *et al. J. biol. Chem.* **260**, 11461-11467 (1985).
9. Sklar, L. A. & Oades, Z. G. *J. biol. Chem.* **260**, 11468-11475 (1985).
10. von Tscharner, V., Derenleau, D. A. & Baggiolini, M. *J. biol. Chem.* **261**, 10163-10168 (1986).
11. Di Virgilio, F., Lew, D. P. & Pozzan, T. *Nature* **310**, 691-693 (1984).
12. Petersen, O. H. & Petersen, C. C. H. *News physiol. Sci.* **1**, 5-8 (1986).
13. Spät, A., Bradford, P. G., McKinney, J. S., Rubin, R. P. & Putney, J. W. Jr *Nature* **319**, 514-516 (1986).
14. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
15. Korchak, H. M., Vienne, K., Rutherford, L. E. & Weissmann, G. *Fedn Proc.* **43**, 2749-2754 (1984).
16. Fukushima, Y. & Hagiwara, S. *J. Physiol., Lond.* **358**, 255-284 (1985).
17. Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752-754 (1981).
18. Yellen, G. *Nature* **296**, 357-359 (1982).
19. Barrowman, M. M., Cockcroft, S. & Gomperts, B. D. *Nature* **319**, 504-507 (1986).
20. Sawyer, D. W., Sullivan, J. A. & Mandell, G. L. *Science* **230**, 663-666 (1985).
21. Bøyum, A. *Scand. J. clin. Invest.* **21** (Suppl. 97), 77-79 (1986).
22. Hesketh, T. R., Smith, G. A., Moore, J. P., Taylor, M. V. & Metcalfe, J. C. *J. biol. Chem.* **25**, 4876-4882 (1983).
23. Almers, W. & Neher, E. *FEBS Lett.* **192**, 13-18 (1985).

A complementary DNA clone for a macrophage-lymphocyte Fc receptor

Victoria A. Lewis, Terry Koch, Helen Plutner & Ira Mellman*

Department of Cell Biology, Yale University School of Medicine,
333 Cedar Street, PO Box 3333, New Haven,
Connecticut 06510, USA

Macrophages, granulocytes and many lymphocytes express or secrete receptors for the Fc domain of immunoglobulins (Ig)[1]. These Fc receptors (FcRs) are heterogeneous and can be distinguished on the basis of their cellular distribution and specificities for different immunoglobulin isotypes. Although their functions are not completely understood, FcRs are known to be involved in triggering various effector cell functions and in regulating differentiation and development of B-cells[1-4]. One of the best characterized is the mouse macrophage-lymphocyte receptor for IgG_1 and IgG_{2b} (ref. 5). On macrophages, this FcR mediates the endocytosis of antibody-antigen complexes via coated pits and coated vesicles[6,7], the phagocytosis of Ig-coated particles[8], and the release of various inflammatory and cytotoxic agents[1]. It is possible that the receptor possesses an intrinsic ligand-activated ion channel activity responsible for some of these functions[9,10]. The IgG_1/IgG_{2b} FcR has been isolated and shown to be a transmembrane glycoprotein of relative molecular mass (M_r) 47,000–60,000 (47–60 K) containing four *N*-linked oligosaccharide chains and a large (>10K) cytoplasmic domain[11]. It is also immunologically indistinguishable from the murine Ly-17 alloantigen which, in turn, is tightly linked to the *Mls* lymphocyte activation locus[12-14]. Here we describe the isolation and characterisation of a complementary DNA clone encoding the whole of the IgG_1/IgG_{2b} FcR expressed by the mouse macrophage-like cell line $P388D_1$. The receptor is a member of the immunoglobulin superfamily and, like Ly-17, maps to the distal portion of chromosome 1. cDNA probes detect one or two mRNA species in FcR^+ macrophage and B-cell lines, but not in FcR^- cells or a receptor-deficient variant derived from a FcR^+ B-cell line[15]. Finally, DNA hybridization analysis indicates the receptor gene is partially deleted or rearranged in the FcR^- variant.

FcR cDNAs were isolated from a $P388D_1$ λgt11 library (obtained from A. Ezekowitz, Harvard University) using both antibody and oligonucleotide probes. Antibody screening was performed[16] using a monospecific rabbit antiserum raised against sodium dodecyl sulphate (SDS)-denatured FcR purified from the J774 mouse macrophage cell line[11]. A single positive phage was plaque-purified and found to contain a 580 base pair (bp) insert (pFcRAb). For oligonucleotide screening, a double stranded 42 bp probe corresponding to residues 2-15 of the FcR's chemically determined N-terminal amino-acid sequence was made (Fig. 1*a*). Two plaques (pFcR11 and pFcR13) were purified, subcloned and found to contain 0.8 and 1.5 kilobase (kb) inserts respectively. The clone isolated by antibody screening cross-hybridized with those obtained by oligonucleotide screening (not shown). Their specificity for the FcR was confirmed by nucleotide sequence analysis of the 5′ ends of pFcR11 and pFcR13. Both contained sequences corresponding to the N-terminal amino acid-sequence (Fig. 1*b*).

Figure 1*c* shows the complete nucleotide sequence compiled from the three inserts, and the amino acid sequence of the FcR thereby deduced. The largest clone, pFcR13, includes the receptor's entire coding region as well as 3′ untranslated sequences containing a consensus $poly(A)^+$ addition sequence and a $poly(A)^+$ tract. pFcR11 corresponds to the 5′ half of the 1.5 kb pFcR13 clone, whereas pFcRAb corresponds to the 3′ half: they overlap with each other by only 27 nucleotides. From the deduced amino acid sequence, the FcR contains an N-terminal hydrophobic signal peptide. However, some ambiguity remains regarding the identity of the actual initiator AUG codon as the putative signal sequence contains three methionine residues (at positions −13, −29 and −39) any of which could begin a signal sequence. We believe that Met−13 is the best candidate because the putative signal it begins contains the hydrophobic and cleavage domains typical of shorter signal peptides[17]. Although Met−29 lies in a nucleotide sequence most closely (but not precisely) corresponding to a consensus start site for translation[18], it implies a signal peptide beginning with an atypically long stretch of polar residues (G. von Heijne, personal communication).

The amino acid sequence of the mature FcR starts with a Thr and contains four potential *N*-linked glycosylation sites (Asn-X-Ser/Thr, X ≠ Pro; at positions 36, 63, 137 and 144), a single hydrophobic membrane-spanning segment followed by a charged stop-transfer sequence, and a large hydrophilic cytoplasmic domain. These results agree well with the results of our previous biochemical investigations which used endoglycosidase digestion to identify four *N*-linked oligosaccharides and protease digestion of ^{35}S-methionine-labelled microsomes to locate a cytoplasmic segment possibly >10 kD in length[11]. While the length of the cytoplasmic domains predicted from the position of the first stop codon is only 47 residues (6.7 K), the size determined from SDS polyacrylamide gel electrophoresis data could easily be an overestimation, as found previously[19]. Nevertheless it is interesting that the first stop codon is followed by two long open reading frames of 32 and 65 amino acids prior to the occurrence of multiple stop sites. Although the digestion experiments also showed that ~15 K of the receptor was resistant to proteolysis from either side of the membrane, the nucleotide sequence reveals that this protection was not due to the presence of multiple membrane-spanning domains. As the FcR contains only a single hydrophobic domain it is probable that its putative ion channel activity[9,10] could occur only by receptor aggregation. However, the FcR's predicted membrane spanning segment of 29 amino acids is somewhat longer than those found in most membrane proteins.

---

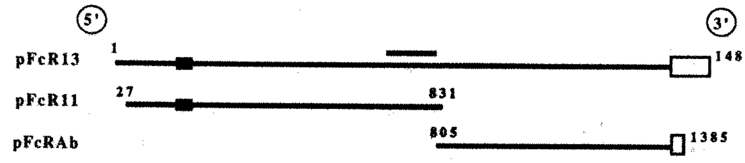
* To whom correspondence should be addressed.

a

```
    2                                      15                   22
   HisAspLeuProLysAlaValValLysLeuGluProProTrp-x-GlnValLeuLysGluAsp...

        *                 *  *
   CATGATCTTCCAAAGGCTGTGGTCAAACTCGAGCCCCCGTGGATCCAGGTGCTCAAGGAAGAC...(pFcR13)
   ||||| || || ||||||||||| || || ||||| || |||
5'-CATGACCTICCIAAGGCTGTGGTGAAGCTIGAGCCICCITGG-3'
3'-GTACTGGAIGGITTCCGACACCACTTCGAICTCGGIGGIACC-5'
```

b



c

```
                              -39                        -30                           -20
                              MetGlyIleLeuProPheLeuLeuIleProMetGluSerAsnTrpThrValHisValPhe
TGTGCCTGCAGCTGACTCGCTCCAGAGCTGATGGGAATCCTGCCGTTCCTACTGATCCCCATGGAGAGCAACTGGACTGTCCATGTGTTC  90
                           -10                        -1 +1                        10
SerArgThrLeuCysHisMetLeuLeuTrpThrAlaValLeuAsnLeuAlaAlaGlyThrHisAspLeuProLysAlaValValLysLeu
TCACGGACTTTGTGCCATATGCTACTGTGGACAGCCGTGCTAAATCTTGCTGCTGGGACTCATGATCTTCCAAAGGCTGTGGTCAAACTC  180
                        20                            30                            40
GluProProTrpIleGlnValLeuLysGluAspThrValThrLeuThrCysGluGlyThrHisAsnProGlyAsnSerSerThrGlnTrp
GAGCCCCCGTGGATCCAGGTGCTCAAGGAAGACACGGTGACACTGACATGCGAAGGGACCCACAACCCTGGGAACTCTTCTACCCAGTGG  270
                        50                            60
PheHisAsnGlyArgSerIleArgSerGlnValGlnAlaSerTyrThrPheLysAlaThrValAsnAspSerGlyGluTyrArgCysGln
TTCCACAATGGGAGGTCCATCCGGAGCCAGGTCCAAGCCAGCTACACGTTTAAGGCCACAGTCAATGACAGTGGAGAATATCGGTGTCAA  360
                        80                            90                            100
MetGluGlnThrArgLeuSerAspProValAspLeuGlyValIleSerAspTrpLeuLeuLeuGlnThrProGlnLeuValPheLeuGlu
ATGGAGCAGACCCGCCTCAGCGACCCTGTAGATCTGGGAGTGATTTCTGACTGGCTGCTGCTCCAGACCCCTCAGCTGGTGTTTCTGGAA  450
                        110                           120                           130
GlyGluThrIleThrLeuArgCysHisSerTrpArgAsnLysLeuLeuAsnArgIleSerPhePheHisAsnGluLysSerValArgTyr
GGGGAAACCATCACGCTAAGGTGCCATAGCTGGAGGAACAAACTACTGAACAGGATCTCGTTCTTCCATAATGAAAAATCCGTGAGGTAT  540
                        140                           150                           160
HisHisTyrSerSerAsnPheSerIleProLysAlaAsnHisSerHisSerGlyAspTyrTyrCysLysGlySerLeuGlyArgThrLeu
CATCACTACAGTAGTAATTTCTCTATCCCAAAAGCCAACCACAGTCACAGTGGGGACTACTACTGCAAAGGAAGTCTAGGAAGGACACTG  630
                        170                           180                           190
HisGlnSerLysProValThrIleThrValGlnGlyProLysSerSerArgSerLeuProValLeuThrIleValAlaAlaValThrGly
CACCAGTCCAAGCCTGTCACCATCACTGTCCAAGGGCCCAAGTCCAGCAGGTCTTTACCAGTATTGACAATTGTGGCTGCTGTCACTGGG  720
                        200                           210                           220
IleAlaValAlaAlaIleValIleIleLeuValSerLeuValTyrLeuLysLysLysGlnValProAspAsnProProAspLeuGluGlu
ATTGCTGTCGCAGCCATTGTTATTATCCTAGTATCCTTGGTCTATCTCAAGAAAAAGCAGGTTCCAGACAATCCTCCTGATCTGGAAGAA  810
                        230                           240                           250
AlaAlaLysThrGluAlaGluAsnThrIleThrTyrSerLeuLeuLysHisProGluAlaLeuAspGluGluThrGluHisAspTyrGln
GCTGCCAAAACTGAGGCTGAGAATACGATCACCTACTCACTTCTCAAGCATCCCGAAGCCCTGGATGAAGAAACAGAGCATGATTACCAG  900
         *              260                           270                           280
AsnHisIleENDSerProLeuAlaLeuGlyLysAlaSerGlnLysGlyGlnAspLeuValSerProGlyProArgAspAlaValAspIle
AACCACATTTAGTCTCCCTTGGCATTGGGAAAAGCAAGCCAGAAAGGCCAGGATCTAGTGTCTCCTGGTCCAAGGGATGCTGTAGATATT  990
                        290                           300                           310
LysGluAsnIleGlnSerHisPheCysGluSerENDAsnGlnGlnThrLeuArgAspTrpPheProMetValAspCysThrAsnAspSer
AAAGAAAACATCCAGAGTCACTTCTGTGAGTCCTGAAACCAACAGACACTACGAGATTGGTTCCCAATGGTTGACTGTACTAATGACTCC 1080
                        320                           330                           340
HisAsnLeuGlnLeuProAsnSerArgLeuPheCysTyrArgSerThrLeuProLeuLysLeuIleAsnLeuLeuProLeuArgAspCys
CATAACTTACAGCTTCCCAACTCAAGACTCTTCTGCTATCGATCCACACTGCCACTAAAATTAATCAACTTACTGCCGTTAAGAGACTGC 1170
                        350
SerAsnValGlyThrAlaTyrAsnSerArgENDAspArgGluAsnAsnENDCysEndAspAlaSerSerArgENDENDEND
AGCAACGTGGGAACTGCTTATAATTCACGCTAAGATAAGGAGAACAATTAATGTTAGGATGCATCATCTCGGTGATGATAAACCAGACTT 1260

GTACCAACCTGTCAAAGATGATGCTTCTTTATAAGGAACTTCAATTTTTTTTTTATTTTTCTATAAAGGTCACATGTTACATTTGAGTTC 1350
TTATTTGAATGTTCTTGGACCATGTGTAACAGGACTGCTATCTGAATAAAATACTAGGTGTTGAAAAAAAAAAAAAAAAAAAAAAAA... 1437
```

d

```
MOPC  104E (1)            Q A V V T Q E S - A L T T S P G E T V T L T C R S S T G A V T I S
N-CAM (180)             K P K I T Y V - E N - K T A M E L E D Q I T L T C E A S G D P I P S -
FcR (1)           T H D L P K A V V K L E P P W I Q V L K E D T V T L T C E G T H N P G N S S

MOPC 104E (33)    N Y A N W V Q Q K P D H L F T G L I G G T N N R A P G V P A R F S G S L I G
N-CAM (212)       - - I T W K T S T R N I S N E E K - - - - - - - - - T L D G R I V V R S H A
FcR (39)          - - T Q W F H N G - - - - - - - - - - - - - - - - - - - - - R S I - R S Q V

MOPC 104E (71)    N K A A L T I T G A Q T E D E A I Y F C A L W Y S N H W V F G G G T K L T V L
N-CAM (239)       R V S S L T L K E I Q Y T D A G E Y V C T A S N T - - - I G Q D S Q A M Y L E
FcR (53)          Q - A S Y T F K - A T V N D S G E Y R C Q M E Q T - - - R L S D P V D L G V I
```

Fig. 1 IgG_1/IgG_{2b} Fc receptor nucleotide sequence and predicted amino acid sequence. *a*, Nucleotide sequence of synthetic oligonucleotide probe based on N-terminal amino-acid sequence of the intact FcR protein isolated from J774 cells. Assignment of residue 1 (Thr) was ambiguous and therefore omitted, as was residue 16 (Ile). Overlapping 27- and 24-mers (underlined) corresponding to amino acids 2–15 from coding and non-coding strands were chemically synthesized. Five deoxyinosine residues were inserted at positions of highest ambiguity. Asterisks denote three positions of mismatch between the sequence of the synthetic probe and the isolated cDNA clone. *b*, Schematic representation of cDNA clones representing the complete FcR coding sequence. pFcRAb and pFcR11 correspond to the 3′ and 5′ halves of the coding sequence respectively and overlap by 27 nucleotides just 3′ to sequences coding for the putative transmembrane domain (overlined). pFcR13 completely contains both pFcRAb and pFcR11. Black boxes, sequences corresponding to the synthetic oligonucleotide probe. Open boxes, poly$(A)^+$ tracts. *c*, Nucleotide and predicted amino-acid sequence of pFcR13. Nucleotides are numbered in the 5′ to 3′ direction starting with the first nucleotide in the codon of the N-terminal residue in the mature protein. Nucleotides 5′ to base 1 have negative numbers and include residues encoding a putative signal peptide and a 5′ untranslated region. The four glycosylation sites are lightly overlined, cysteine residues involved in forming Ig-like domains are boxed, and the consensus polyadenylation signal is underlined. The single putative transmembrane domain is indicated by the heavy overline, and is followed by a hydrophilic stop-transfer sequence of three lysine residues. *d*, Amino-acid sequence alignment of FcR, N-CAM, and a mouse λ-light chain (MOPC 104E) Ig-like domains. Common residues are boxed, and gaps inserted by dashes to achieve maximum alignment. The amino-acid residue that begins each line is given in parentheses. FcR sequence corresponds to the first Ig domain; sequence of the MOPC 104E variable region is given. Similar alignments can be obtained using rabbit polymeric IgR, murine Thy-1, Ia antigen β-chain, or Lyt-2 sequences. We thank Alan F. Williams (MRC, Oxford) for assistance with this analysis.

Methods. *a*, FcR was purified to homogeneity by immunoaffinity chromatography using an anti-FcR monoclonal antibody, 5F2a[11], coupled to a 0.5 ml epoxy-activated, silcon-based HPLC matrix provided by Beckman Instruments (final concentration 30 mg IgG per ml resin). The detergent phase of a TX-114 lysate was prepared from 5 ml packed J774 cells (as previously described for immunoprecipitation[11]), filtered and passed once over the column at 0.2 ml min^{-1} at 4 °C. Following extensive washing with Tris-buffered saline, bound material was eluted with 0.5 M propionic acid containing 25 mM β-D-octylglucoside, immediately neutralized and dialysed exhaustively against glass-distilled water. Final yield was 39 μg of FcR per 2.5×10^9 cells. N-terminal sequence analysis was performed on a modified Beckman 890C spinning cup sequenator and allowed the unambiguous assignment of 20 contiguous amino acids with gaps at positions 1 and 16. *b*, A $P388D_1$ λgt-11 cDNA library was screened using a monospecific polyclonal anti-FcR antiserum followed by ^{125}I-labelled protein A essentially as described[16]. A single clone was isolated and designated pFcRAb. The library was then re-screened using a double-stranded, high-specific-activity oligonucleotide probe as described by Ullrich *et al.*[25]. All λgt-11 isolates were plaque-purified and subcloned into pUC9 or Bluescribe plasmid vectors. *c*, DNA sequencing was performed on both strands of all three cDNA clones by the method of Sanger *et al.*[26] after insertion into the M13 bacteriophage vector mp19. Sequential overlapping subclones were produced by the method of Dale *et al.*[27]. pFcRAb was also partially sequenced (on only one strand) by the method of Maxam and Gilbert[28].

The amino acid sequence of the IgG_1/IgG_{2b} FcR bears striking similarities to immunoglobulin and certain immunoglobulin-like molecules. The receptor's external domain contains four evenly spaced Cys residues (positions 28, 70, 109, 153) which define an internal repeat. These repeats have conserved amino-acid sequences surrounding each Cys which are typical of V-related immunoglobulin-like domains[20] suggesting the presence of interchain disulphide bonds forming two such domains. Particularly good matches exist between sequences adjacent to Cys 28, Cys 70 of FcR and rabbit polymetic immunoglobulin receptor[19] domain I (not shown), immunoglobulin V_L-domains such as MOPC 104E, and N-CAM[21] (Fig. 1*d*). Thus the patterns around the Cys residues seem closest to V and V-like domains but the length of sequence between the Cys residues is far too short to account for a V-like fold. If β-strands are assigned on the basis of sequence similarities then it seems possible that the FcR domains may be folded in the C-domain pattern (A. F. Williams, personal communication). Computer searches of nucleic acid and protein sequence databases (GenBank, Staden) indicated further similarities between the FcR and other immunoglobulin-like molecules including class II MHC antigen β-chain, Lyt-2 antigen and murine Thy-1 (which, like N-CAM,

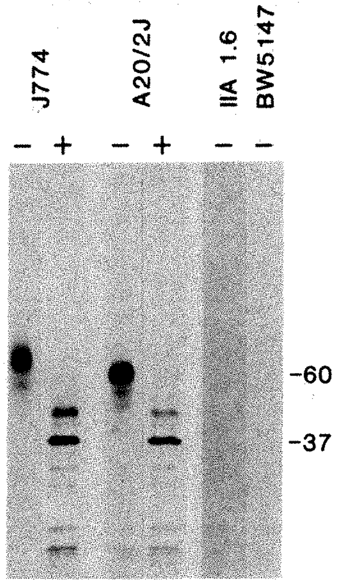


Fig. 2 Immunoprecipitation of ^{35}S-FcR from macrophage, B- and T-lymphoma cell lines. J774, A20/2J, IIA-1.6 and BW5147 cells were metabolically labelled with ^{35}S-methionine, lysed using Triton X-114 and immunoprecipitated with anti-FcR antiserum as described previously[11]. Radiolabelled bands corresponding to the mature receptor were detected in lysates of J774 cells (64K) and A20/2J cells (60K). Endoglycosidase F treatment (+ lanes) of immunoprecipitated ^{35}S-FcR gave identical 37K peptides (arrow), and several minor bands resulting from incomplete digestion or proteolytic degradation[11]. Both the T-lymphoma BW5147 and the selected variant of A20/2J, IIA-1.6, failed to synthesize immunoprecipitable FcR. Molecular weights in kilodaltons.

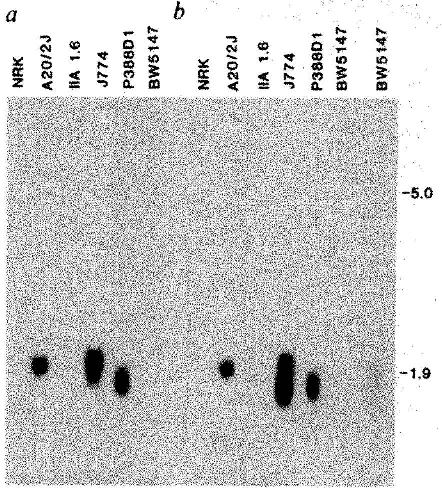


Fig. 3 Northern blot analysis of poly(A)$^+$ RNA isolated from FcR-positive and -negative cell lines. RNAs were prepared by guanidine-HCl lysis and phenol/chloroform extraction. Poly(A)$^+$RNA was selected by oligo-dT cellulose chromatography, resolved on 1.0% agarose-formaldehyde gels, and transferred to Zetabind nylon membranes. RNAs were hybridized with the 3′ encoding clone, pFcRAb, *a*, stripped at 95 °C with 0.1 × SSC/0.1% SDS, and re-hybridized using the 5′ encoding clone, pFcR11, *b*. Purified inserts were labelled by nick-translation to a specific activity of 0.5×10^9 c.p.m. per μg. Hybridization and washings (0.2 × SSC/0.1% SDS at 50 °C) were carried out essentially as described[29]. Cell lines used included rat fibroblasts (NRK, FcR$^-$), A20/2J B-cells (FcR$^+$), the FcR$^-$ A20 variant IIA-1.6, the macrophage-like cell lines J774 and P388D$_1$ (both FcR$^+$), and the T-cell lymphoma BW5147 (FcR$^-$). Autoradiographs represent 5 h exposures, except for BW5147 which was exposed for 5 h and 48 h (panel B). The 3′ probe pFcRAb hybridized with single bands only in FcR$^+$ cell lines (1.9 kb in A20/2J and J774, 1.7 kb in P388D$_1$) (panel A). The 5′ probe pFcR11 recognized an additional lower molecular weight band (1.6 kb) in J774 cells and in BW5147 cells after long exposures (using the 3′ probe, no hybridization was detected after similarly long exposures). Sizes of murine 28S (5.0 kb) and 18S (1.95 kb) ribosomal RNAs are indicated.

is expressed in neural tissues)[21].

As mentioned above, the IgG_1/IgG_{2b} FcR has been shown to be immunologically related to the Ly-17 alloantigen, a polymorphic locus which maps to mouse chromosome 1 (refs 12, 13). Thus, it was of interest to determine the chromosomal localization of the gene recognized by FcR cDNA. Using labelled pFcRAb, we probed DNA from a chromsome panel derived from CHO × mouse fibroblast somatic cell hybrids (provided by D. Pravtcheva and F. Ruddle)[22]. Conditions were chosen for an *Eco*RI digestion such that pFcRAb hybridized with a single murine band that was easily differentiated from is CHO homologue (Fig. 4). Analysis of DNA from 9 hybrid cell lines consistently assigned the FcR gene to mouse chromosome 1 (not shown). We also probed DNA from a second CHO × mouse panel (provided by C. Kozak) that included cell lines characterized by having undergone translocations involving only portions of mouse chromosomes (ref. 23). These results indicate that the FcR structural gene, and therefore the Ly-17 structural gene, is uniquely localized to the region of chromosome 1 distal to the *Pep*-3 locus[14].

To confirm that the cDNAs detected mRNA only in FcR$^+$ cells, and to investigate the relationship between receptor species expressed in different cell types, hybridization analysis was performed using poly(A)$^+$ RNA isolated from FcR$^+$ or FcR$^-$ cells. Because the presence or absence of FcR is usually determined only on the basis of surface expression[1], we first analysed cell lines by immunoprecipitation after labelling with ^{35}S-methionine. Using a rabbit anti-FcR antiserum, labelled receptor was precipitated from two macrophage cell lines (J774 and P388D$_1$, only the former is shown) and one B-cell line (A20/2J) (Fig. 2). Although the apparent relative molecular masses of the macrophage and B-cell products were different (64K against 60K), the difference was due to glycosylation heterogeneity, as Endo F treatment of the two immunoprecipitates yielded a 37K deglycosylated peptide. Two cell lines, T-cell line BW5147 and a variant of the A20/2J B-cell line, IIA-1.6 (obtained from B. Jones), selected *in vitro* for the loss of surface FcR[15] were shown not to synthesize immunoprecipitable FcR (Fig. 2). Expression of other surface antigens (such as surface immunoglobulin and Ia) in the IIA-1.6 variant, however, is the same as its normal parent A20/2J[15].

Northern blot analysis using pFcRAb as a probe only detected single transcripts in mRNA from the FcR$^+$ cell lines (Fig. 3*a*). The J774 macrophage and A20 B-cell lines contained mRNAs of similar sizes (1.9 kb) whereas the P388D$_1$ transcript appeared slightly smaller (1.7 kb). Somewhat different results were obtained using pFcR11 as a probe. Although pFcR11 hybridized with all the transcripts detected by pFcRAb, an additional band (1.6 kb) was detected in J774 mRNA (Fig. 3*b*). Because pFcR11 corresponds to the 5′ half of the FcR cDNA and pFcRAb corresponds to the 3′ half (Fig. 1*b*), these results suggest that some cell lines (for example, J774) may make two related FcR species which are more homologous at their N-terminal than at their C-terminal regions. We are now determining whether multiple FcR isoforms are indeed expressed.

Although the 5′ probe (pFcR11) detected small amounts of the 1.9 and 1.6 kb transcripts in BW5147 cells, neither the 3′ probe (pFcRAb) nor the 5′ probe hybridized with poly(A)$^+$ RNA from fibroblasts or from the FcR$^-$ A20/2J variant, IIA-1.6 (Fig. 3). To investigate the mechanism underlying the loss of FcR-specific mRNA from the variant cell line, a preliminary Southern blot analysis was performed. Using the 3′-end probe pFcRAb, single bands were detected in J774 and A20/2J DNA digested with *Eco*RI (15 kb), *Hin*dIII (10.5 kb), or *Xba*I (6.8 kb)

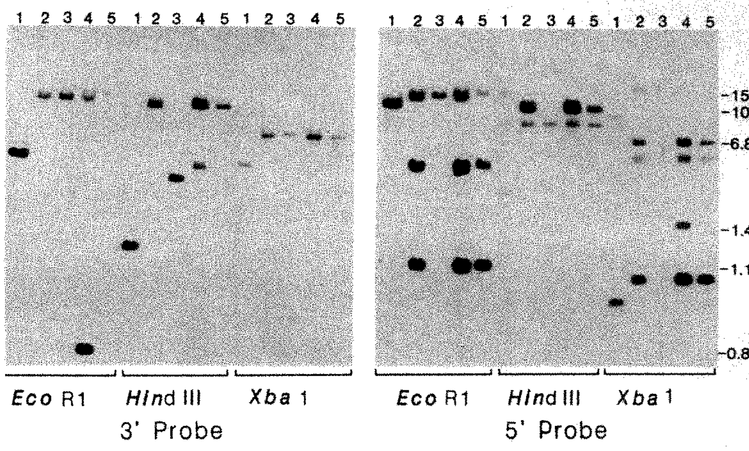


Fig. 4 Southern blot analysis of FcR-positive and -negative celllines. DNA isolated from: track 1, CHO cells (FcR$^-$); track 2, A20/2J; track 3, IIA-1.6; track 4, BW5147; and track 5, J774 cells was digested to completion with *Eco*RI, *Hin*dIII or *Xba*I, separated on a 0.8% agarose gel (10 μg per track) and transferred to nylon (Zetabind) membrane. The blots were then hybridized consecutively with the 3′ probe pFcRAb (left panel) or the 5′ probe pFcR11 (right panel) essentially as described[29]. Isolated inserts were labelled by nick-translation, filters washed with 0.2×SSC, 0.1% SDS at 55 °C, and the autoradiographs exposed overnight. Sizes of the major restriction fragments are shown in kb.

(Fig. 4, left panel). In contrast, *Hin*dIII digestion of IIA-1.6 DNA produced a single hybridizable fragment (4 kb), much smaller than that found in the A20/2J parent (10.5 kb). *Hin*dIII and *Eco*RI digests of the FcR$^-$ BW5147 DNA produced additional smaller bands (5.0 kb and 0.8 kb, respectively) not seen in FcR$^+$ cell lines. When the same blots were hybridized with the 5′-end probe pFcR11, 1-3 extra bands were detected in all cell lines except the receptor-negative variant IIA-1.6 (Fig. 4, right panel). In IIA-1.6 DNA, the 5′ probe even failed to detect the major 4 kb *Hin*dIII and 6.8 kb *Xba*I fragments seen with the 3′ probe (Fig. 4, right panel). As before, pFcR11 also hybridized with a unique low molecular weight band in BW5147 DNA. These data strongly suggest that the failure of the IIA-1.6 variant, and possibly BW5147, to express hybridizable FcR mRNA is attributable to deletions or rearrangements of the FcR gene.

Because the IgG_1/IgG_{2b} receptor studied here is only one of a large family of isotype- and cell type-specific FcRs[1], the existence of multiple FcR-related genes is not surprising. The availability of FcR cDNA clones and cell lines containing alterations in the receptor gene will now permit analysis of the distribution, structure and activities of these receptors, and the definitive establishment of their relationship to functionally-defined genetic loci of immunological importance, such as *Mls*[12-14,24].

We thank Adrian Hayday, Elisabetta Ullu, Chris Tschudi, Arthur Horwich, Saumyen Sarkar, Johanna Griffin, Christine Howe, Susan Froschauer and Heini Miettinen for their advice and assistance; Gary Davis and David Speicher for protein sequencing; Alan Williams and Gunnar von Heijne for help with sequence analysis; Christine Kozak, Dimi Prartcheva and Frank Ruddle for providing DNA from their chromosome panels; Alan Ezekowitz and Brigitte Huber for providing cDNA libraries; Barry Jones for providing the IIA-1.6 cell line and Pam Ossorio for photographic work. Supported by NIH grants and a Presidental Research Award from The Leukemia Society of America, Inc.

Note added in proof: During preparation of this manuscript, we learned that Ravage *et al.* (*Science*, in the press) have also cloned and sequenced FcR specific cDNAs. The sequence they obtained from a clone designated β-2 (isolated from a J774 macrophage cell line cDNA library) is identical to that described here (P388D1-derived cDNA). Although identical, the two sequences arise from mouse strains representing different *Mls* alleles (BALB/c: *Mls* b and DBA/2: *Mls* a, respectively). Therefore it is clear that the FcR cloned here, although immunologically identical to the Ly-17 alloantigen, cannot be the product of the *Mls* locus.

Received 14 August; accepted 24 September 1986.

1. Unkeless, J. C., Fleit, H. & Mellman, I. *Adv. Immun.* **31,** 247-270 (1981).
2. Phillips, N. E. & Parker, D. C. *J. Immun.* **132,** 627-632 (1984).
3. Bijsterbosch, M. K. & Klaus, G. G. B. *J. exp. Med.* **162,** 1825-1836 (1985).
4. Daeron, M., Yodoi, J., Neauport-Sautes, C., Moncuit, J. and Fridman, W. H. *Eur. J. Immun.* **15,** 662-667 (1985).
5. Mellman, I. & Unkeless, J. C. *J. exp. Med.* **152,** 1048-1069 (1980).
6. Ukkonen, P., Lewis, V. A., Marsh, M., Helenius, A. & Mellman, I. *J. exp. Med.* **163,** 952-971 (1986).
7. Mellman, I. & Plutner, H. *J. Cell Biol.* **98,** 1170-1176 (1984).
8. Mellman, I., Plutner, H., Steinman, R. M., Unkeless, J. C. & Cohn, Z. A. *J. Cell Biol.* **96,** 887-895 (1983).
9. Young, J. D.-E., Unkeless, J. C., Young, T. M., Mauro, A. & Cohn, Z. A. *Nature* **306,** 186-189 (1983).
10. Young, J. D.-E., Unkeless, J. C., Kaback, H. R. & Cohn, Z. A. *Proc. natn. Acad. Sci. U.S.A.* **80,** 1636-1640 (1983).
11. Green, S. A., Plutner, H. & Mellman, I. *J. Biol. Chem.* **260,** 9867-9874 (1985).
12. Holmes, K. L., Palfree, R. G. E., Hämmerling, U. & Morse, H. C., III *Proc. natn. Acad. Sci. U.S.A.* **82,** 7706-7710 (1985).
13. Hibbs, M. L., Hogarth, P. M. & McKenzie, F. C. *Immunogenetics* **22,** 335-348 (1985).
14. Kozak, C. A., Davidson, W. F. & Morse, H. C., III *Immunogenetics* **19,** 163-168 (1984).
15. Jones, B., Tite, J. P. & Janeway, C. A. Jr *J. Immun.* **136,** 348-356 (1986).
16. Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80,** 1194-1198 (1983).
17. von Heijne, G. *J. molec. Biol.* **173,** 243-251 (1984).
18. Kozak, M. *Nucleic Acids Res.* **12,** 857-872 (1984).
19. Mostov, K. E., Freidlander, M. & Blobel, G. *Nature* **308,** 37-43 (1984).
20. Williams, A. F., Barclay, A. N., Clark, M. J. & Gagnon, J. in *Gene expression during normal and malignant differentiation.* (eds Andersson, L. C., Gahmberg, C. G. & Ekblom, P.) (Academic, 1985).
21. Hemperly, J. J., Murray, B. A., Edelman, G. M. & Cunningham, B. A. *Proc. natn. Acad. Sci. U.S.A.* **83,** 3037-3041 (1986).
22. D'Eustachio, P., Pravtcheva, D., Marcu, K. & Ruddle, F. H. *J. exp. Med.* **151,** 1545-1550 (1980).
23. MacInnes, J. I., Morris, V. L., Flintoff, W. F. & Kozak, C. A. *Virology* **132,** 12-25 (1984).
24. Festenstein, H. *Transplant. Proceed.* **8,** 339-342 (1976).
25. Ullrich, A. *et al.* *Nature* **309,** 418-425 (1984).
26. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463-5467 (1977).
27. Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmid* **13,** 31-40 (1985).
28. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65,** 499-560 (1980).
29. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).

Loss of mouse fibroblast cell response to phorbol esters restored by microinjected protein kinase C

Gabriella Pasti*, Juan-Carlos Lacal†,
Barbour S. Warren*, Stuart A. Aaronson†
& Peter M. Blumberg*

* Molecular Mechanisms of Tumor Promotion Section,
Laboratory of Cellular Carcinogenesis and Tumor Promotion, and
† Laboratory of Cellular and Molecular Biology, National Cancer
Institute, Bethesda, Maryland 20892, USA

The phorbol esters in addition to being among the most potent mouse skin tumour promoters profoundly affect many different biological systems[1,2]. It is postulated that they act through activation of protein kinase C (ref. 3-5), but substantial heterogeneity in their pharmacological and binding behaviour in some systems[6] has caused concern about whether this is their only target. Evidence linking protein kinase C activation with biological responses to the phorbol esters includes similarity in structure-activity relations for binding and response[7]; *in vitro* phosphorylation of specific proteins by protein kinase C at the same sites at which phorbol ester treatment induces phosphorylation in intact cells[3]; and corre-

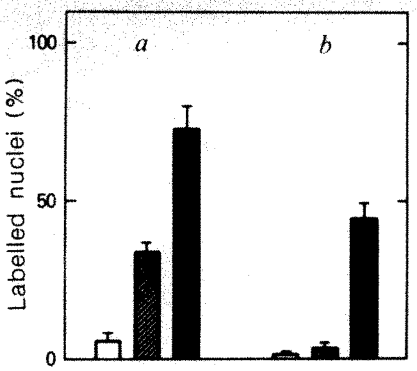


Fig. 1 Blocking of the mitogenic response of Swiss 3T3 cells to phorbol esters by pretreatment with phorbol ester. Control cells (*a*) or cells previously incubated with phorbol esters (*b*) were treated with no addition (□), 200 nM PDBu (▨) or 10% calf serum (■), and DNA synthesis measured by autoradiography. Data represent the mean ± s.e.m. for 5 experiments, with 200 cells counted for each experimental condition in each experiment.

Methods. Swiss 3T3 cells were plated at 1.2×10^5 cells per 35-mm tissue culture dish in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% calf serum (CS, Hyclone) and incubated at 37 °C in 95% air: 5% CO$_2$. Two days after plating, the cells were refed with fresh DMEM containing 10% CS. After an additional 2 days, the plates were washed 5× with DMEM, and DMEM containing 1% CS was added back in the presence or absence of 200 nM PDBu. Three days later all the cultures were washed 5 times with 1 mg ml^{-1} bovine serum albumin (BSA) in phosphate-buffered saline to remove residual PDBu. Conditioned DMEM containing 1% CS was added back together with 2 μCi ml^{-1} ^{3}H-thymidine (NEN) and 200 nM PDBu or 10% CS. After incubation for a further 24 h, the plates were washed and subjected to autoradiography as described[15,16], and labelled nuclei counted.

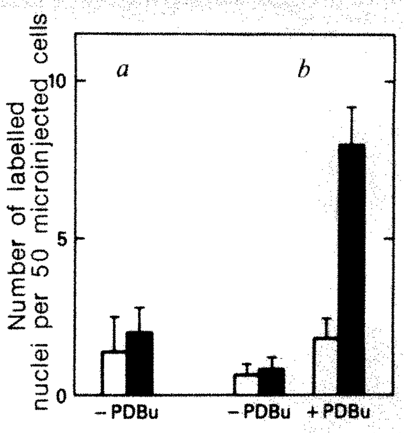


Fig. 2 Stimulation of mitogenesis by microinjected protein kinase C in phorbol-ester pretreated cells. Control cells (*a*) or cells pretreated with phorbol esters (*b*) were microinjected with protein kinase C(■) or with carrier solution (□). They were then incubated with ^{3}H-thymidine for 24 h in the absence or presence of 200 nM PDBu, and labelled nuclei in the microinjected area were counted. Data represent the mean ±s.e.m. for 5 experiments using two separate preparations of protein kinase C. In each experiment, 50–100 cells were microinjected under each experimental condition.

Methods. Swiss 3T3 cells were grown in 60-mm tissue-culture dishes and treated as described (Fig. 1 legend). Protein kinase C was purified to homogeneity as described by Jeng *et al.*[17]. After the final purification step the enzyme was concentrated in the presence of fatty acid-free BSA (Sigma), and the buffer was changed to 130 mM KCl. The final BSA concentration was 15 mg ml^{-1}. The concentrations of protein kinase C in the two preparations used were 2.2 and 4.1 mg ml^{-1}, respectively. The specific activities were 1.6 and 1.2 μmol ^{32}P transferred mg^{-1} min^{-1}, respectively. Activity in the absence of Ca^{2+} and phospholipid was <1% of that in its presence. The enzyme was stored for no more than 3 days at 0 °C before use. During this period, activity decreased no more than 20%. Volumes of 5×10^{-14} l were usually microinjected[16] using a Leitz micromanipulator coupled to a Zeiss IM35 microscope under a F-Achromat LD 32/0.4 PH1 lens (320× total). Pipettes were made from borosilicate glass tubes of 0.7–1 mm outer diameter (Kimble) using a vertical pipette puller Model 720 (D. Kopf Instruments, Tujinga, California). Treatment of tips and loading of pipettes by suction from the tips was as described previously[16,18]. Labelling of microinjected control cells incubated in the presence of PDBu was not determined because the results would have been obscured by the mitogenic response of uninjected cells within the microinjected area.

lation in certain cell types between down regulation of protein kinase C on chronic phorbol ester treatment and loss of cellular responsiveness to the phorbol ester[8–10]. Here we report that microinjection of purified protein kinase C into Swiss 3T3 fibroblasts pretreated with the phorbol ester phorbol 12,13-dibutyrate (PDBu) restores the mitogenic response of the cells to PDBu, directly establishing the involvement of protein kinase C in this response.

Quiescent cultures of Swiss 3T3 cells were obtained in the presence of low serum (Fig. 1 legend). Under these conditions addition of 200 nM PDBu stimulates DNA synthesis, measured by autoradiography, in 34 ± 3% (mean ± s.e.m.) of the cells compared with 83 ± 7% (mean ± s.e.m.) of cells stimulated with 10% serum (Fig. 1*a*). Pretreatment of the cultures for 3 days with 200nM PDBu markedly reduces the number of cells responding to PDBu (Fig. 1*b*). The number of cells responding to 10% serum was also reduced although by a smaller percentage perhaps because the phorbol ester pretreatment resulted in higher cell density (protein per plate = 144% of controls).

Similar results were obtained when DNA synthesis was measured by total ^{3}H-thymidine incorporation. In control cells, stimulation of ^{3}H-thymidine incorporation by PDBu was 35 ± 6% (mean ± s.e.m.) of that by 10% serum (5 experiments). In PDBu pretreated cells PDBu caused no stimulation, whereas 10% serum stimulated ^{3}H-thymidine incorporation to 78 ± 12% (mean ± s.e.m.) of the level for cells that were not pretreated.

To determine the effect of restoration of protein kinase C on the mitogenic response, cells were microinjected with $4\text{–}8\times10^5$ molecules per cell of active protein kinase C. This amount was comparable to the initial kinase levels in control cells of $3\text{–}6\times10^5$ (ref. 8). To control for effects caused by microinjection, results were compared with those for cells microinjected with carrier protein only. Microinjected protein kinase C mitogenically stimulates phorbol ester pretreated cells provided 200 nM PDBu was present (Fig. 2). No stimulation was seen in the absence of PDBu in either control or pretreated cells, confirming that the mitogenic effect is phorbol ester-dependent. The absolute level of mitogenic stimulation, 17% of nuclei labelled, was lower than that observed for stimulation of control cells by PDBu. This difference may be more apparent than real, however, because these results are normalized to the total number of microinjected cells. Swiss 3T3 cells are somewhat fragile to the microinjection procedure under our culture conditions and ~40% of the cells lysed after injection either with carrier protein or with enzyme. The number of labelled nuclei per surviving microinjected cell would therefore be larger.

These studies provide direct evidence for the role of protein kinase C in the mitogenic response of Swiss 3T3 cells to the phorbol esters. The role of protein kinase C in mitogenesis is of particular interest because there is evidence that some growth factors and oncogene products induce enhanced phosphatidylinositol turnover, activating protein kinase C (ref. 11), although apparently protein kinase C cannot account for all their mitogenic activity[10]. Our results do not support the suggestion of Collins and Rozengurt[10], based on the kinetics of recovery of response in the presence of cycloheximide, that phorbol ester pretreatment imposes an additional postreceptor block to responsiveness, but this difference may reflect the difference in stimulation conditions.

A similar strategy may be applicable to other systems, for example, cell-cell communication[12], in which desensitization is observed in response to chronic phorbol ester treatment. Microinjection has been used to demonstrate a role for protein kinase C in two other systems: enhancement of calcium current in neurones of *Aplysia*[13]; and K$^+$ ion conductance in photoreceptors of *Hermissenda*[14].

We thank Dr Stuart H. Yuspa for critical reading of the manuscript.

Received June 10; accepted 22 August 1986.

1. Slaga, T. J. (ed.) *Mechanisms of Tumor Promotion* (Raven, New York, 1984).
2. Blumberg, P. M. *Crit. Rev. Toxicol.* **8,** 153–234 (1981).
3. Nishizuka, Y. *Nature* **308,** 693–698 (1984).
4. Niedel, J. E., Kuhn, L. J. & Vandenbark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **80,** 36–40 (1983).
5. Leach, K. L., James, M. & Blumberg, P. M. *Proc. natn. Acad. Sci. U.S.A.* **80,** 4208–4212 (1983).
6. Blumberg, P. M. *et al. Biochem. Pharmac.* **33,** 933–940 (1984).
7. Blumberg, P. M. *et al. Mechanisms of Tumor Promotion* Vol. 3 (ed. Slaga, T. J.) 143–184 (CRC, Boca Raton, Florida, 1984).
8. Collins, M. K. L. & Rozengurt, E. *J. cell. Physiol.* **112,** 42–50 (1982).
9. Rodriguez-Pena, A. & Rozengurt, E. *Biochem. biophys. Res. Commun.* **120,** 1053–1059 (1984).
10. Collins, M. K. L. & Rozengurt, E. *J. cell. Physiol.* **118,** 133–142 (1984).
11. Berridge, M. J. *Biotechnology* 541–546 (1984).
12. Enomoto, T. & Yamaski, H. *Cancer Res.* **45,** 3706–3710 (1985).
13. DeRiemer, S. A., Strong, J. A., Albert, K. A., Greengard, P. & Kaczmarek, L. K. *Nature* **313,** 313–316 (1985).
14. Farley, J. & Auerbach S. *Nature* **319,** 220–223 (1986).
15. Adams, R. L. P. *Expl Cell Res.* **56,** 49–54 (1969).
16. Lacal, J. C., Srivastava, S. K., Anderson, P. S. & Aaronson, S. A. *Cell* **44,** 609–617 (1986).
17. Jeng, A. Y., Sharkey, N. A. & Blumberg, P. M. *Cancer Res.* **46,** 1966–1971 (1986).
18. Graessmann, A., Graessmann, M. & Müeller, C. *Meth. Enzym.* **65,** 816–825 (1980).

The v-*fms* oncogene induces factor independence and tumorigenicity in CSF-1 dependent macrophage cell line

Esther F. Wheeler, Carl W. Rettenmier,
A. Thomas Look & Charles J. Sherr

Department of Tumor Cell Biology, St Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, Tennessee 38105 USA

The McDonough strain of feline sarcoma virus (SM-FeSV) transforms fibroblast cell lines in culture and produces fibrosarcomas in domestic cats[1,2]. SM-FeSV does not induce haematopoietic malignancies in spite of the fact that its viral oncogene, v-*fms*, codes for a glycoprotein related to the receptor for the mononuclear phagocyte colony stimulating factor, CSF-1 (refs 3, 4). The v-*fms*-coded polypeptide includes the complete extracellular domain of the c-*fms* proto-oncogene product[5,6] and retains the ability to bind CSF-1 specifically[4]. The two molecules have very similar sequences except at their extreme carboxyl terminal ends where 40 amino acids of the c-*fms*-coded glycoprotein are replaced by 11 unrelated residues in the v-*fms* product[5]. Autophosphorylation of the c-*fms* gene product on tyrosine is enhanced by CSF-1 addition[3], whereas phosphorylation of the v-*fms*-coded glycoprotein appears to be constitutive[4]. We now show that introduction of the v-*fms* gene into simian virus-40 (SV40)-immortalized, CSF-1 dependent macrophages renders them independent of CSF-1 for growth and tumourigenic in nude mice. These factor-independent cell lines express unaltered levels of the c-*fms* product which is down-modulated in response to either CSF-1 or the tumour promoter 12-*O*-tetradecanoyl-phorbol-13-acetate (TPA). The induction of factor independence by a non-autocrine mechanism suggests that the v-*fms* product is an unregulated kinase that provides growth stimulatory signals in the absence of ligand.

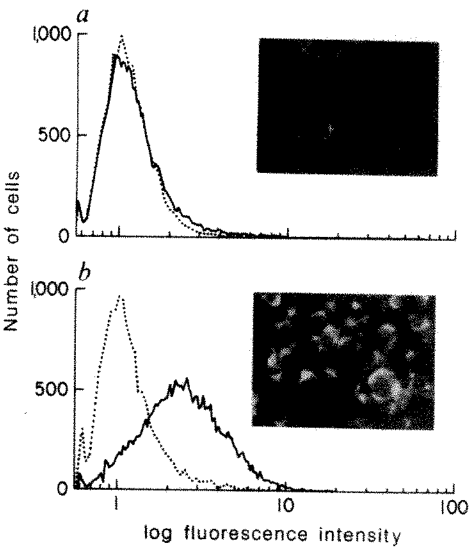


Fig. 1 Expression of v-*fms*-coded antigens in SM-FeSV-infected BAC1.2F5 macrophages. SM-FeSV-infected cells maintained for 5 passages in *a*, the presence or *b*, the absence of CSF-1 were tested for expression of v-*fms*-coded epitopes identified by the rat monoclonal antibody SM2.6.3[11,13]. Solid line, the flow cytometric fluorescence profile recorded after binding of SM2.6.3 to the surface of viable cells; dotted line, the control fluorescence profile obtained with an irrelevant isotype-matched rat myeloma protein. The insets in each panel show photomicrographs (400 × magnification) of fixed, permeabilized cells incubated with the same monoclonal antibody and subsequently reacted with an affinity purified, fluorescein-labelled goat antibody to rat immunoglobulin G (IgG). Note the single positive cell in the field of factor dependent cells shown in *a*. The frequency of positive cells in these cultures was <2% of the total population (see Table 1). BAC1.2F5 cells studied in parallel showed background fluorescence, similar to that observed for the majority of cells in *a*.

Methods. Indirect immunofluorescence labelling of viable cells and flow cytometric analysis were performed as described[11], except that cells scraped from culture dishes were preincubated in medium containing 100 mg ml^{-1} bovine IgG before staining, to prevent nonspecific binding of the monoclonal antibody to Fc receptors. Immunofluorescence staining of methanol-fixed cells was as described[13], except that the slides were pretreated with buffer containing 100 mg ml^{-1} of bovine IgG before reaction with the monoclonal antibody. The insets in *a* and *b* are matched exposures of SM-FeSV-infected FD and FI cells, respectively, photographed in the same experiment.

The murine macrophage cell line BAC1 was derived from adherent BALB/c × A.CA F$_1$ mouse spleen cells by transfection with origin-defective SV40 DNA[7]. The 2F5 subclone[8], chosen for its inability to grow in the absence of CSF-1, produces lysozyme, collagenase and esterase, secretes interleukin-1, expresses Fc receptors, and is capable of Fc-mediated phagocytosis. BAC1.2F5 cells were infected with high titre, helper-free SM-FeSV pseudotype virus[6] prepared by transfection of cloned, replication-defective SM-FeSV DNA[9] into the packaging NIH-3T3 cell line, Ψ2, which contains a defective Moloney murine leukaemia provirus that provides viral replicative functions but cannot package its own RNA[10]. No factor-independent colonies were obtained from mock-infected BAC1.2F5 cells. SM-FeSV-infected cells grown for seven days in the absence of CSF-1 gave ~0.1% factor-independent (FI) colonies. These were grown and maintained in the absence of CSF-1. Virus-infected factor-dependent (FD) cell lines were maintained in L-cell conditioned medium containing CSF-1 and passaged at 5-day intervals.

SM-FeSV-infected macrophages were plated in semisolid medium with or without CSF-1. Uninfected BAC1.2F5 cells formed colonies in agar only in the presence of CSF-1 (Table 1); FI cells formed colonies (6.9%) in the absence of CSF-1; as did FD cells at passage 3 and 5 after infection, but at low frequency (0.1–0.3%). The plating efficiency of FD cells was greatly increased (6.0%) after 14 passages. When factor-

Table 1 Efficiency of colony formation of SM-FeSV-infected BAC1.2F5 cells

	Colony forming efficiency (% of plated cells) +CSF-1	−CSF-1
Uninfected BAC1.2F5 cells	7.0	<0.01
FD cells:		
passage 3	6.8	0.1
passage 5	4.6	0.3
passage 14	10.9	6.0
FI cells:		
passage 5	9.3	6.9
passage 10	13.4	4.2

A plasmid containing biologically active SM-FeSV proviral DNA[9] was transfected into Ψ2 cells[10], and transformed Ψ2 subclones were derived from single cells by endpoint dilution in microtitre culture dishes. Supernatants from a subclone producing 2×10^7 helper-free SM-FeSV focus-forming units (FFU) per ml on NIH-3T3 cells[6] were used to infect 3×10^5 BAC1.2F5 cells at 10 FFU per cell. Cells were grown in Dulbecco's minimal medium containing 15% fetal calf serum and 25% murine L-cell conditioned medium containing CSF-1. Two days after infection the culture was divided and infected cells were continuously maintained in the presence or absence of CSF-1. In the absence of CSF-1 most of the infected cells failed to survive, but about 0.1% of the unpassaged cells gave rise to factor independent colonies which were grown up and maintained in the absence of CSF-1 (FI cells). Parallel cultures maintained in medium containing the growth factor (FD cells) gave rise to FI cells at a progressively increased frequency when shifted to factor-free medium. Cells from both infected populations were plated at various passage levels after infection in 60 mm culture dishes in 0.3% Noble agar containing complete medium supplemented with or lacking CSF-1. In each experiment duplicate sets of cultures were established containing 10^3, 10^4 and 10^5 cells per dish. The efficiency of colony formation was independent of the cell density and represents an average of triplicate determinations. Continuously growing colonies (>100 cells) were scored three weeks after plating. Colonies picked from SM-FeSV-infected cells plated in the absence of CSF-1 gave rise to cell lines which were indistinguishable in their properties from the original FI cell line derived in liquid medium.

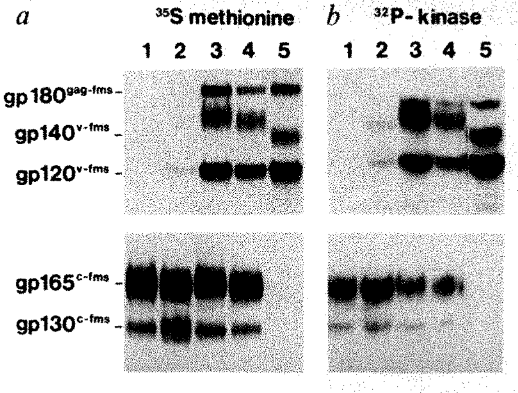


Fig. 2 Synthesis (*a*) and tyrosine kinase activities (*b*) of v-*fms* (top panels) and c-*fms* (bottom panels) gene products in infected mouse cells. The lanes are the same in each panel. Lane 1, uninfected BAC1.2F5 cells; lane 2, SM-FeSV-infected FD cells maintained in medium containing CSF-1 for 5 passages; lane 3, infected FD cells maintained for 14 passages; lane 4, SM-FeSV-infected FI cells at passage 5; lane 5, SM-FeSV-transformed mouse NIH-3T3 fibroblasts. The mobilities of the various forms of v-*fms* and c-*fms* gene products are noted at the left margins. Due to differences in oligosaccharide processing, the mature form of the v-*fms*-coded glycoprotein (gp140$^{v\text{-}fms}$) in the mouse macrophages has a greater apparent molecular mass than that expressed in NIH-3T3 cells. Matched exposures of the corresponding autoradiograms are shown in upper and lower panels.

Methods. To avoid down modulation of CSF-1 receptors[18] in cells grown in the presence of CSF-1, all cell lines were shifted to factor-free medium 18 h before the experiment. *a*, Cells were metabolically radiolabelled with ^{35}S-methionine for 90 min and then lysed in RIPA buffer (50 mM Tris HCl, *p*H 7.4, containing 150 mM NaCl, 20 mM EDTA, 1% Triton X-100, 1% sodium deoxycholate and 0.1% sodium dodecyl sulphate (SDS)) with 2% Aprotinin and 1 mM phenylmethylsulphonyl fluoride as protease inhibitors. Unlabelled cell lysates for protein kinase assays were prepared in parallel. The lysates were incubated first with antibodies that react with the v-*fms*-coded glycoproteins but do not bind the product of murine c-*fms* gene[13]. After recovery of the immune complexes using *Staphylococcus aureus* protein-A-Sepharose as immunoadsorbant, the cleared lysates were incubated with a rabbit antiserum to a recombinant v-*fms*-coded polypeptide expressed in bacteria that cross reacts with the murine CSF-1 receptor[3,14]. *b*, Unlabelled immune complexes were incubated in the presence of [γ^{32}P]ATP and divalent cation as described[16,17]. Radiolabelled products in washed immunoprecipitates were separated by SDS gel electrophoresis and detected by fluorography (*a*) or autoradiography (*b*).

independent agar colonies from FD cells at passages 5 and 14 were picked and cultured in liquid medium without CSF-1, the resulting subclones were indistinguishable from unpassaged FI cells directly after infection. Thus SM-FeSV-infected cells maintained in medium containing CSF-1 progressively gave rise to FI cells. Because the SM-FeSV was produced from the Ψ2 packaging cell line, this was not due to virus spread. No RNA-dependent DNA polymerase or focus-forming activity was found in the culture fluids, so no virus was produced by infected cultures.

From Southern blot analysis with a v-*fms* probe, early passage FD cells contained an average of 0.5–1.0 proviral DNA copies per cell, integrated polyclonally in host DNA, but late-passage FD cells, and FI subclones derived from single cells contained multiple (~2–4) proviruses per cell (data not shown). Most FD cells must have acquired proviral DNA after infection, but only those with multiple proviral insertions appeared to have a proliferative advantage, even in the presence of CSF-1.

Conditioned medium from FI cultures was tested using uninfected BAC1.2F5 cells to exclude the possibility that FI cells produced CSF-1. BAC1.2F5 cells die within 7 days in the absence of CSF-1. No CSF-1 activity was detected, so the proliferation of FI cells is not due to an autocrine mechanism. Northern blot analysis using a CSF-1 complementary DNA probe confirmed that BAC1.2F5, FD and FI cells did not express CSF-1 messenger RNA.

In fibroblasts infected with SM-FeSV the v-*fms*-coded glycoprotein, gp140$^{v\text{-}fms}$, is required at high levels on the cell surface for transformation[11,12]. Only a low percentage of unpassaged SM-FeSV-infected FD cells gave rise to FI colonies when CSF-1 was removed from the medium, suggesting that most of these cells produced insufficient amounts of the glycoprotein to become factor independent. We analysed the infected cultures by fluorescence-activated flow cytometry using a monoclonal antibody to the extracellular, amino-terminal domain of the v-*fms*-coded polypeptide[11,13]. FD cells at passage 5, like uninfected BAC1.2F5 cells, did not express the v-*fms*-coded glycoprotein (Fig. 1*a*). In contrast, FI cells at passage 5 expressed relatively high levels of the glycoprotein (Fig. 1*b*). The FI cells were uniformly fluorescence-positive when assayed for cytoplasmic fluorescence (Fig. 1*b*, inset), whereas less than 2% of FD cells were fluorescent (Fig. 1*a*, inset). Thus, only a small proportion of early passage SM-FeSV-infected BAC1.2F5 cells express sufficient v-*fms* gene product to become factor independent.

We also examined the expression of the feline v-*fms* and murine c-*fms* gene products in the various cell lines. Cultures were metabolically radiolabelled with ^{35}S-methionine, lysed with detergent and sequentially immunoprecipitated; the first antibody was specific for the v-*fms* gene product[13] (Fig. 2*a*, top), and the second, a rabbit antiserum to a recombinant v-*fms*-coded polypeptide, cross-reacts with the murine CSF-1 receptor[3,14] (Fig. 2*a*, bottom). Because v-*fms* sequences were inserted into the open reading frame of the SM-FeSV *gag* gene, the primary

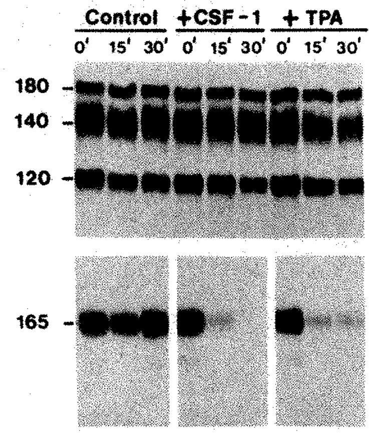


Fig. 3 Down modulation of the c-*fms* gene product in SM-FeSV-infected FI cells. Upper panel, v-*fms* gene products; lower panel, c-*fms* gene products. FI cells grown in factor-free medium were metabolically labelled for 15 min with ^{35}S-methionine and incubated for an additional 75 min in medium containing a 100-fold excess of unlabelled methionine. Following the 'chase' period, parallel tissue culture dishes were incubated in normal medium (control) or in medium containing either 3 nM CSF-1 or 1×10^{-7} M TPA for the indicated periods of time (0, 15, 30 min). Cells were lysed in RIPA buffer. The clarified lysates were incubated first with antibodies that react only with the v-*fms*-coded glycoproteins, and then with an antiserum to the c-*fms* product as in Fig. 2 legend. The immunoprecipitates were analysed by SDS gel electrophoresis. The sizes of the expected glycoproteins are given in kilodaltons (left). The figure gives data from quantitative sequential immunoprecipitations and shows matched exposures of the corresponding autoradiograms.

translation product is a glycosylated polyprotein of relative molecular mass (M_r) 180,000 (180K), (gP180$^{gag-fms}$) which is proteolytically cleaved, yielding an immature form of the v-*fms*-coded glycoprotein (gp120^{v-fms}), which undergoes processing and appears at the cell surface as a glycoprotein of greater M_r, gp140^{v-fms} (refs 13, 15). Uninfected BAC1.2F5 cells did not express v-*fms* gene products (Fig. 2*a*, top, lane 1). SM-FeSV-infected FD cells at passage 5 produced low levels of these glycoproteins (lane 2), but after serial subculturing in the presence of CSF-1, these cells (passage 14; lane 3) expressed comparable levels of the viral oncogene products to FI cells (lane 4) and SM-FeSV-transformed NIH-3T3 cells (lane 5).

Uninfected and infected BAC1.2F5 cells, whether CSF-1 dependent or not, expressed high levels of the murine c-*fms* gene product (Fig. 2*a*, bottom, lanes 1-4). The glycoprotein gp130^{c-fms} is an intracellular precursor of the mature 165K cell surface receptor, gp165^{c-fms} (ref. 14). SM-FeSV transformed NIH/3T3 cells did not express the c-*fms* product (lane 5). Immune complex protein kinase assays[16,17] show that the v-*fms* (Fig. 2*b*, top) and c-*fms* (Fig. 2*b*, bottom) glycoproteins were phosphorylated in ratios proportional to the expression levels detected by metabolic radiolabelling. Phosphoamino acid analysis confirmed that *in vitro* phosphorylation was exclusively on tyrosyl residues. Subculturing of SM-FeSV-infected FD cells yields a population expressing high levels of the v-*fms*-coded glycoproteins without affecting synthesis of the c-*fms* product.

Since both the v-*fms* and c-*fms* products were coexpressed in FI cells, we studied down-modulation of the c-*fms* gene product[18] after addition of CSF-1 or TPA. Cells labelled with ^{35}S-methionine were exposed to CSF-1 or TPA for brief periods, and the v-*fms* and c-*fms* products were immunoprecipitated from detergent lysates and analysed on gels. Turnover of gp140^{v-fms} was unaffected by CSF-1 or TPA, whereas turnover of gp165^{c-fms} was greatly increased by either compound (Fig. 3). Thus, both CSF-1 and TPA down-modulate gp165^{c-fms} but not the transforming glycoprotein.

Since the SM-FeSV-infected FI cells did not produce CSF-1 or show constitutive down-modulation of their CSF-1 receptors, their conversion to factor independence was directly mediated by the v-*fms*-coded receptor kinase rather than an autocrine mechanism. Thus, although the v-*fms* gene product includes the ligand-binding domain of the CSF-1 receptor, it can function as an unregulated enzyme without ligand. This has not been formally shown in fibroblasts, since fibroblast cell lines susceptible to transformation by SM-FeSV produce CSF-1 (ref. 4). The structural differences between the carboxyl termini of the v-*fms* and c-*fms* products could be responsible for the transforming activity of the oncogene product[5]. The carboxyl terminus of the v-*fms*-coded glycoprotein lacks a distal tyrosine residue found in the c-*fms* product, so it may be that phosphorylation of the c-*fms* product on this tyrosine decreases kinase activity, as observed[19] for pp60^{c-src}.

Late passage FD and FI cells both formed tumours when inoculated subcutaneously or intraperitoneally into nude mice (5×10^6 cells per animal), whereas uninfected SV40-immortalized BAC1.2F5 cells did not. Tumours appeared within 5 weeks after injection, spread haematogenously, and formed histiocytic sarcomas metastatic to liver, spleen, pancreas, lung and spinal cord. CSF-1 independent cell lines from tumour tissues had the histological appearance of macrophages, stained cytochemically for butyrate esterase, contained multiple copies of the integrated SM-FeSV provirus and expressed high levels of v-*fms*-coded glycoprotein at the cell surface. The v-*fms* receptor kinase therefore acts with a nuclear oncogene product to form tumours *in vivo*. Critical alterations in the kinase domain of the c-*fms* proto-oncogene might similarly contribute to haematopoietic malignancy.

Oncogenes encoding other tyrosine kinases, including v-*abl*, v-*src*, v-*fps*/*fes* and v-*erb*-B, can transform a variety of different cells of the myeloid and lymphoid lineages[20-27]. The normal physiological functions of the v-*abl*, v-*src*, and v-*fps*/*fes* gene products remain obscure, but the ability of v-*erb*-B (a truncated EGF receptor gene)[28,29] to transform erythroid progenitors suggests that viral tyrosine kinases can generate abnormal signals for cell proliferation. This raises the possibility that the v-*fms* gene product may convert other haematopoietic cell lines to factor independence, including those that do not express CSF-1 receptors.

We thank Dr E. Richard Stanley for BAC1.2F5 cells, Drs Ernest Kawasaki and Peter Ralph of Cetus Corporation for a CSF-1 cDNA clone and Shawn Kramer and Jean Elmendorf for technical assistance. This work was supported by grant CA-38187 (to C.J.S.), by Cancer Center grant CA-21765 from the National Cancer Institute and by the American Lebanese Syrian Associated Charities of St Jude Children's Research Hospital. E.F.W. was supported by a National Research Service Award (CA-09346) and C.W.R. by a Biomedical Research Support grant (RR-05584-21) from the NIH.

Received 7 July; accepted 22 September 1986.

1. McDonough, S. K., Larsen, S., Brodey, R. S., Stock, N. D. & Hardy, W. D. Jr *Cancer Res.* **31,** 953-956 (1971).
2. Sarma, P. S., Sharar, A. L. & McDonough, S. *Proc. Soc. exp. Biol. Med.* **140,** 1365-1368 (1972).
3. Sherr, C. J. *et al. Cell* **41,** 665-676 (1985).
4. Sacca, R., Stanley, E. R., Sherr, C. J. & Rettenmier, C. W. *Proc. natn. Acad. Sci. U.S.A.* **83,** 3331-3335 (1986).
5. Coussens, L. *et al. Nature* **320,** 277-280 (1986).
6. Wheeler, E. F. *et al. J. Virol.* **59,** 224-233 (1986).
7. Schwarzbaum, S., Halpern, R. & Diamond, B. *J. Immun.* **132,** 1158-1162 (1984).
8. Morgan, C. J. & Stanley, E. R. *Biochem. biophys. Res. Commun.* **119,** 35-41 (1984).
9. Donner, L., Fedele, L. A., Garon, C. F., Anderson, S. J. & Sherr, C. J. *J. Virol.* **41,** 489-500 (1982).
10. Mann, R., Mulligan, R. C. & Baltimore, D. *Cell* **33,** 153-159 (1983).
11. Roussel, M. F., Rettenmier, C. W., Look, A. T. & Sherr, C. J. *Molec. cell. Biol.* **4,** 1999-2009 (1984).
12. Nichols, E. J., Manger, R., Hakomori, S., Herscovics, A. & Rohrschneider, L. R. *Molec. cell. Biol.* **5,** 3467-3475 (1985).
13. Anderson, S. J., Furth, M., Wolff, L., Ruscetti, S. K. & Sherr, C. J. *J. Virol.* **44,** 696-702 (1982).
14. Furman, W. L. *et al. Virology* **152,** 432-445 (1986).
15. Anderson, S. J., Gonda, M. A., Rettenmier, C. W. & Sherr, C. J. *J. Virol.* **51,** 730-741 (1984).

16. Barbacid, M. & Lauver, A. V. *J. Virol.* **40,** 812–821 (1981).
17. Rettenmier, C. W., Chen, J. H., Roussel, M. F. & Sherr, C. J. *Science* **228,** 320–322 (1985).
18. Guilbert, L. J. & Stanley, E. R. *J. biol. Chem.* **261,** 4024–4032 (1986).
19. Cooper, J. A., Gould, K. L., Cartwright, C. A. & Hunter, T. *Science* **231,** 1431–1434 (1986).
20. Rosenberg, N. & Baltimore, D. *J. exp. Med.* **143,** 1453–1463 (1976).
21. Pierce, J. H. & Aaronson, S. A. *J. Virol.* **46,** 993–1002 (1983).
22. Adkins, B., Leutz, A. & Graf, T. *Cell* **39,** 439–445 (1984).
23. Kahn, P., Adkins, B., Beug, H. & Graf, T. *Proc. natn. Acad. Sci. U.S.A.* **81,** 7122–7126 (1984).
24. Oliff, A., Agranovsky, O., McKinney, M. D., Murty, V. V. V. S. & Bauchwitz, R. *Proc. natn. Acad. Sci. U.S.A.* **82,** 3306–3310 (1985).
25. Cook, W. D., Metcalf, D., Nicola, N. A., Burgess, A. W. & Walker, F. *Cell* **41,** 677–683 (1985).
26. Pierce, J. H. *et al. Cell* **41,** 685–693 (1985).
27. Anderson, S. M., Klinken, S. P. & Hankins, W. D. *Molec. cell. Biol.* **5,** 3369–3375 (1985).
28. Downward, J. *et al. Nature* **307,** 521–527 (1984).
29. Ullrich, A. *et al. Nature* **309,** 418–425 (1984).

Recurrent mutations in haemophilia A give evidence for CpG mutation hotspots

Hagop Youssoufian*, Haig H. Kazazian Jr*, Deborah G. Phillips*, Sophia Aronis†, George Tsiftis†, Valerie A. Brown‡ & Stylianos E. Antonarakis*

* Genetics Unit, Department of Pediatrics,
Johns Hopkins University School of Medicine,
Baltimore, Maryland 21205, USA
† Coagulation Laboratory, 1st Department of Pediatrics,
Athens University School of Medicine, Athens, Greece
‡ Collaborative Research, Inc., Lexington, Massachusetts 02173, USA

Haemophilia A is a common disorder of blood coagulation caused by a deficiency of factor VIII[1]. It is inherited as an X-linked recessive trait, and one-third of all cases are thought to result from *de novo* mutations[2]. The clinical severity of haemophilia A varies markedly among different families and a subset of the patients with severe disease develop antibodies against factor VIII, called inhibitors[3]. Because of this heterogeneity, it is likely that many different molecular lesions result in haemophilia A. Indeed, of the nine mutations described to date, all appear to be unique changes[4–6]. However in this study of 83 patients with haemophilia A we have identified two different point mutations, one in exon 18 and one in exon 22, that have recurred independently in unrelated families. Each mutation produces a nonsense codon by a change of CG to TG, and each occurred *de novo* on the X-chromosome donated by the maternal grandfather. These observations strongly support the view that CpG dinucleotides are mutation hotspots.

Recent reports have described the isolation and partial sequence of the factor VIII gene and its expression *in vitro* from recombinant DNA clones[7–10]. The gene is 186 kilobases (kb) long and is divided into 26 exons, encoding a protein of 2,351 amino acids. The availability of cloned factor VIII gene fragments has made possible the study of molecular defects in patients with haemophilia A. Because of the clinical heterogeneity of the disorder, the large size of the factor VIII gene, and the pool of new mutations, it has been suggested that the number of genetic lesions producing haemophilia A may be very large[4,5].

We have analysed the genomic DNA of 83 unselected patients with haemophilia A who were referred to us for carrier detection and prenatal diagnosis. Our screening strategy for factor VIII mutations involved restriction analysis with *Taq*I, *Sst*I and *Eco*RI, and successive hybridizations with two complementary DNA probes that span exons 1–12 (probe A) and 14–26 (probe BC) of the factor VIII gene. We detected 6 deletions and 10 point mutations using this strategy (ref. 5 and manuscript in preparation); 9 of the 10 point mutations were detected with *Taq*I, which has the dinucleotide CpG in its recognition sequence.

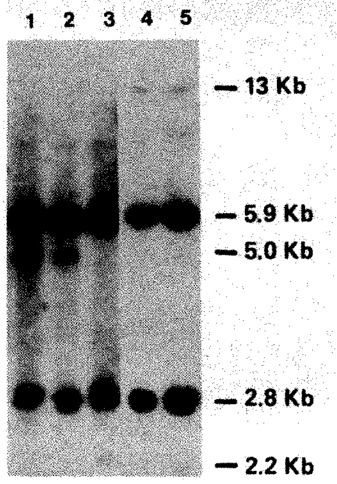


Fig. 1 Autoradiogram of DNAs from members of family K and family T digested with *Taq*I and hybridized to probe BC. Family K: lane 1, subject III-1; lane 2, subject II-2; lane 3, subject I-2; family T: lane 4, subject III-2; lane 5, subject II-2.

Methods. Genomic DNA was isolated from 10–15 ml of peripheral blood by standard techniques. DNA (5 μg) was digested to completion with *Taq*I. Gel electrophoresis, transfer to nitrocellulose filters, and hybridization with a ^{32}P-labelled cloned factor VIII DNA fragment (probe BC) were performed as previously described[5]. The following cloned factor VIII DNA fragments were used as probes: (1) probe A, a 1.7 kb *Kpn*I cDNA fragment that spans exons 1–12; and (2) probe BC, a 6.5 kb *Eco*RI cDNA fragment that spans exons 14–26. These probes were kindly provided by Drs J. Toole and J. Wozney of Genetics Institute, Inc. The 2.2 kb fragments of subject II-2 (family K) and subject III-2 (family T) are difficult to see on this autoradiogram. Sizes of fragments are given in kb.

In an earlier study from our laboratory[5], a patient with severe disease (factor VIII titre <1 U per 100 ml) and factor VIII inhibitors was shown to have a point mutation in exon 18 of the factor VIII gene, where a change of C to T in a CpG dinucleotide led to the generation of a nonsense codon. That family was of Anglo-Saxon origin, and the mutation was present in at least four generations. Upon screening the DNAs of our more recent patients, we observed a *Taq*I site alteration in an unrelated Greek family (family K) where the two affected brothers (subjects III-1 and III-2, Fig. 2*a*) had severe haemophilia A (factor VIII titre <1 U per 100 ml) and were inhibitor-negative. This change appeared identical to the mutation described previously.

Figure 1 shows a representative restriction analysis of genomic DNA from various members of this family digested with *Taq*I. Hybridization with probe BC revealed a new 5.0 kb fragment and absence of a 2.2 kb fragment in an affected male (subject III-1). The DNA of the mother, an obligate heterozygote (Fig. 2*a*, subject II-2), contained the 5.0 kb fragment as well as the 2.2 kb fragment. The patient's father, maternal grandfather and maternal grandmother (Fig. 2*a*; subjects II-1, I-1, and I-2, respectively) displayed the normal *Taq*I pattern, indicating that this mutation appeared *de novo* in one of the mother's X-chromosomes. No variant fragments were detected when the DNA was digested with *Sst*I, *Eco*RI, or *Bam*HI, or when probe A was used for hybridization (data not shown).

To determine the precise nature of the mutation responsible for the altered *Taq*I site, we used oligonucleotides to detect single base mismatches[11,12]. A pair of oligonucleotides were made, each 18 nucleotides long, one having the normal sequence flanking the exon 18 *Taq*I site, the other differing from the normal by the substitution of a T for C in the *Taq*I site, the presumed mutant sequence. Digestion of genomic DNA with *Hin*dIII normally produces a 2.7 kb fragment which contains

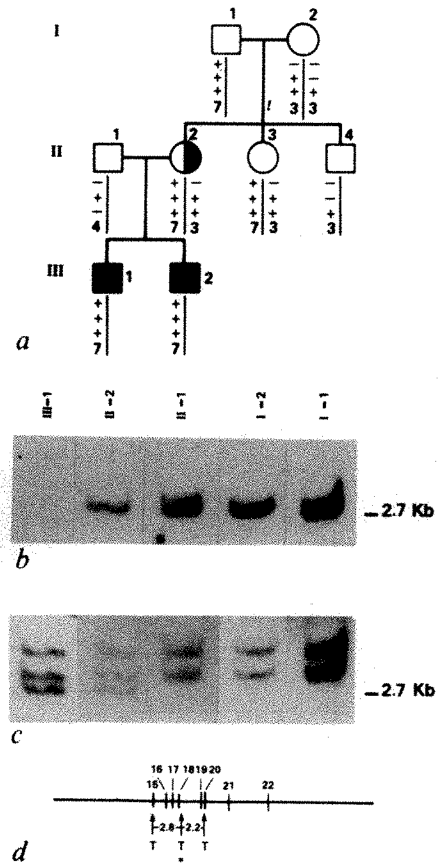


Fig. 2 Analysis of the factor VIII exon 18 mutation. *a*, Pedigree of family K (□, normal male; ■, haemophilic male; ○, normal female; ◑, carrier female). Haplotypes of polymorphic restriction sites are indicated below each subject, in the order (top to bottom) *Bcl*I-factor VIII, *Bgl*I-factor VIII, *Bgl*II-DXS15, *Taq*I-DXS52. The presence or absence of a polymorphic restriction site is shown as + or −. In one case (DXS52) the alternative alleles are arbitrarily numbered. *b*, Oligonucleotide analysis of genomic DNA from members of family K digested with *Hin*dIII and hybridized to the normal exon 18 probe (GATTCGATGGTATCTGCT). *c*, Hybridization of the same gel with an oligonucleotide corresponding to the mutant exon 18 probe (GATTTGATGGTATCTGCT). *d*, Partial restriction map of factor VIII DNA cleaved with *Taq*I (T). Asterisk, location of the absent *Taq*I site in exon 18.

Methods. DXS15 cloned DNA was kindly provided by Dr K. Davies and DXS52 cloned DNA by Dr J. L. Mandel. Oligonucleotides (18 bases long) were chemically synthesized by Dr C. Riley of the Johns Hopkins University School of Medicine, end-labelled using T_4-polynucleotide kinase and ^{32}P-ATP, further purified by electrophoresis on a 20% polyacrylamide sequencing gel, and the labelled probe extracted by the crush and soak method[11,12]. Hybridization to dried agarose gel was performed at calculated melting temperature (T_m minus 4 °C for 3 h, washed in 2× SSPE/0.1% SDS, then washed further in 2×SSPE at 2 °C below T_m for 2 min), and autoradiographed. The agarose gel was re-used after stripping off the first probe by denaturation in 0.5 M NaOH/0.15 M NaCl for 30 min, and neutralization.

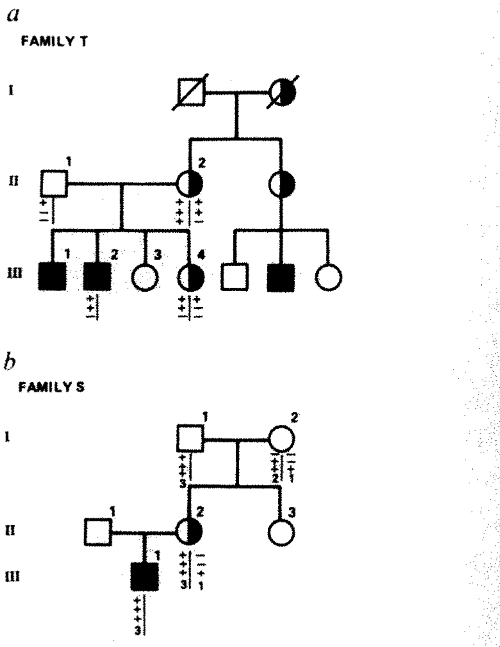


Fig. 3 Pedigrees of *a*, family T; *b*, family S. Both families show the factor VIII exon 22 mutation. Symbols as for Fig. 2*a*.

exon 18 of the factor VIII gene. The normal probe did not hybridize to the 2.7 kb fragment in *Hin*dIII-cleaved DNA from the affected male (Fig. 2*b*, subject III-1), although it did hybridize to the DNAs of unaffected family members. By contrast, the mutant probe hybridized to the 2.7 kb factor VIII fragment from the DNA of the affected male (subject III-1) and his mother (subject II-2), but not to that of the normal members of the family (Fig. 2*c*). Constant hybridizing sequences that were not factor VIII were seen in all individuals. We have thus shown that the basis of the exon 18 *Taq*I site mutation is a C to T change producing a nonsense codon, TGA, at position 1,960 in place of the normal CGA, which codes for arginine. The factor VIII protein produced by this mutant gene is presumably a truncated, inactive molecule.

To determine the parental origin of the new mutation in the patient's mother, we performed DNA polymorphism analysis on various family members. The DNA polymorphisms used were: *Bcl*I and *Bgl*I sites in the factor VIII gene; a *Bgl*II site adjacent to DXS15; and *Taq*I polymorphic alleles detected by probe DXS52[5,13-15]. Analysis of these sites suggested that the X-chromosome of the patient was inherited from the maternal grandfather (Fig. 2*a*), and, therefore, the mutation to an abnormal factor VIII gene occurred in a germ cell of the maternal grandfather. The grandfather was 32 years old at the time his carrier daughter was conceived.

A different *Taq*I site alteration was found in both of two other unrelated families. The haemophilic subjects from each family have severe disease (factor VIII titres <1 U per 100 ml) but are inhibitor-negative. *Taq*I digestion of genomic DNA from subject III-2 of family T (Fig. 1) and subject III-1 of family S (data not shown) and hybridization to probe BC showed an abnormal 13 kb fragment and the absence of the normal 5.9 kb fragment that includes exon 22 sequences[4]. Female carriers of the mutation from family T (subjects III-4 and II-2, Fig. 3*a*) and subject II-2 from family S (Fig. 3*b*) also contained the 13 kb fragment in their DNA, but normal males from either family did not. Neither the maternal grandfather nor the maternal grandmother (subjects I-1 and I-2, respectively) of family S contained the abnormal fragment in their DNA, indicating that the mutation arose *de novo* within two generations.

We determined the precise nucleotide change by making a pair of oligonucleotides, each 18 nucleotides long, flanking the exon 22 *Taq*I site which differed in a single substitution of a T for C in the *Taq*I site. *Hin*dIII digestion of genomic DNA normally produces a 13 kb fragment which contains exon 22 sequences. The normal oligonucleotide probe failed to hybridize to the 13 kb fragment from the affected males (Fig. 4*a*) of family T (III-2) or family S (III-1), but the oligonucleotide bearing the mutant sequence showed specific hybridization to DNA of each affected individual (Fig. 4*b*) as well as to female carriers from each family (subjects II-2 and III-3 of family T and subject II-2 of family S). Therefore, the exon 22 mutation in both these

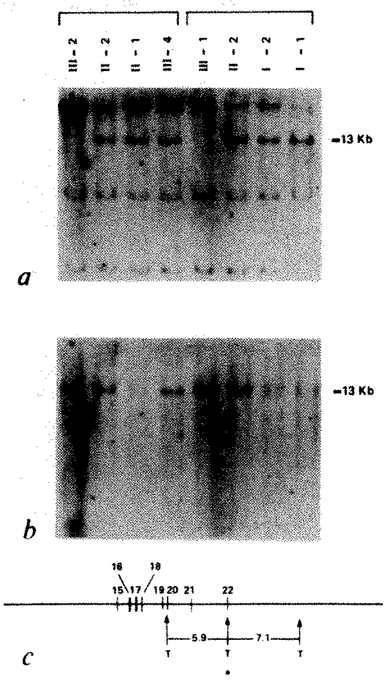


Fig. 4 Analysis of the factor VIII exon 22 mutations by oligonucleotide hybridization to *Hin*dIII-cleaved DNA from members of family T and family S. *a*, Hybridization with an 18-base long oligonucleotide corresponding to the normal exon 22 sequence (CTTATÇGAGGAAATTCCA). *b*, Hybridization of the same gel with an oligonucleotide corresponding to the mutant exon 22 sequence (CTTATŢGAGGAAATTCCA). *c*, Partial restriction map of factor DNA cleaved with *Taq*I (T); asterisk, location of the absent *Taq*I site in exon 22. See Fig. 2 (legend) for methods.

families is due to a CGA to TGA change at codon 2,135 of the factor VIII gene, generating a nonsense codon.

By analysis of polymorphic restriction sites in the factor VIII gene and flanking it, we showed that the *de novo* mutation in family S arose on the X-chromosome of the maternal grandfather (subject I-1, Fig. 3*b*) who was 34 years old at the time his carrier daughter was conceived. It is interesting that both *de novo* mutations were associated with advanced paternal age; further data are needed to determine the significance of this observation.

In all four families, the haplotype of the two intragenic polymorphic sites (*Bcl*I and *Bgl*I) happened to be the same for chromosomes bearing the mutant gene. To confirm that these mutations were independent in origin and to exclude the possibility that the stated parentage was not correct in either family K or family S, we examined multiple, highly polymorphic loci that lie outside the factor VIII gene[15,16]. We examined four such loci; in no case did the mothers in family K (Fig. 2*a*, subject II-2) or in family S (Fig. 3*b*, subject II-2) show a DNA fragment which was not consistent with stated parentage. The probability of the stated parentage being correct is estimated for each patient's mother as 1×10^5 to 1.

These cases are two instances of recurrent mutation in man. The previously reported presence of a particular mutation in two very different ethnic groups and in two or more different β-globin gene frameworks provides strong circumstantial evidence for recurrent mutation in various β-globin gene alleles, including β^S, β^E, and certain β-thalassaemia alleles[17]. Chehab *et al.* have also presented strong evidence for recurrence of a common Mediterranean β-thalassaemia allele[18]. However, unlike β-thalassaemia, there is no positive selection for mutations that produce haemophilia A and the life of each mutation is estimated to be two to four generations in a stationary population[2]. Thus, without positive selection for such mutations, each new mutation would be expected to last 100-150 years. This means that every unrelated affected family should have a different mutant allele (barring recurrent mutation) and that study of these alleles should give an unbiased observation of mutation hotspots.

There is strong circumstantial evidence that CpG dinucleotides are hotspots for mutation in man. The proportion of CpG in sequenced genes is 37% of that expected[19]. Several authors have postulated that methylated cytosines could be mutation hotspots because 5-methylcytosine can spontaneously deaminate to thymine resulting in a C-T transition in DNA[20-22]. Barker *et al.*[23] found an unusually high frequency of polymorphism at a number of loci in human DNA using restriction enzymes whose recognition sequences contain a CpG dinucleotide. Significantly, of the four point mutations producing haemophilia A reported before this study, three were CG-TG changes and one a CG-CA mutation[4-6]. The latter would result from a C-T transition at the CpG on the anti-sense strand of DNA. Our observation of recurrent CG-TG mutations strongly supports the view that these dinucleotides are mutation hotspots. One can estimate the number of recurrences of the codon 1,960 or codon 2,135 mutations in human history given that (1) their recurrence frequency is comparable to that observed in our studies (about 1 in every 50 affected males); (2) the incidence of haemophilia A is 1 in 10,000 males in all racial groups; and (3) each mutation persists for an average of 100-150 years in the population. Using these assumptions, we estimate that over 1,000 individuals are alive today with different recurrent origins of the C-T mutation at codon 1,960 or of the C-T mutation at codon 2,135 of the factor VIII gene, and that the total number of recurrences of each type of mutation in man is perhaps 10,000. Even if this is an overestimate by a factor of 10, it is in striking contrast to recurrence frequencies of 5-10 predicted for the β-globin alleles mentioned earlier, β^S, β^E, and certain β-thalassaemia mutations[17].

We thank Drs J. Toole and J. Wozney for the factor VIII clones, K. Davies for the DXS15 clone, J. L. Mandel for clone DXS52, Drs A. Chakravarti and H. Donis-Keller for helpful discussions, Ms C. E. Dowling for technical assistance and Ms E. Pasterfield for assistance in the preparation of this manuscript. This study was supported by NIH grants to S.E.A. and H.H.K. and a Johns Hopkins Institutional Grant to S.E.A.

Received 6 June; accepted 12 September 1986.

1. Hougie, C. in *Haematology* 3rd edn (eds Williams, W. J., Beutler, E., Erslev, A. J. & Lichtman, M. A.) 1381-1399 (McGraw-Hill, New York, 1983).
2. Haldane, J. B. S. *J. Genet.* **31,** 317-326 (1935).
3. Gill, F. M. in *Factor VIII Inhibitors* (ed. Hoyer, L. W.) 19-29 (Liss, New York, 1984).
4. Gitschier, J. *et al. Nature* **315,** 427-430 (1985).
5. Antonarakis, S. E. *et al. New Engl. J. Med.* **313,** 842-848 (1985).
6. Gitschier, J. *et al. Science* **232,** 1415-1416 (1986).
7. Gitschier, J. *et al. Nature* **312,** 326-330 (1984).
8. Wood, W. I. *et al. Nature* **312,** 330-337 (1984).
9. Vehar, G. A. *et al. Nature* **312,** 337-342 (1984).
10. Toole, J. J. *et al. Nature* **312,** 342-347 (1984).
11. Wallace, R. B. *et al. Nucleic Acids Res.* **9,** 3647-3656 (1981).
12. Studencki, A. B. & Wallace, R. B. *DNA* **3,** 7-15 (1984).
13. Antonarakis, S. E. *et al. Lancet* **i,** 1407-1409 (1985).
14. Goodfellow, P. N., Davies, K. E. & Ropers, H. H. *Cytogenet. Cell Genet.* **40,** 296-352 (1985).
15. Schumm, J. *et al. Cytogenet. Cell Genet.* **40,** 739 (abstract) (1985).
16. Kazazian, H. H. Jr *et al. Am. J. hum. Genet.* **38,** 860 (1986).
17. Orkin, S. H. & Kazazian, H. H. Jr *A. Rev. Genet.* **18,** 131-171 (1984).
18. Chehab, F. F., Honig, G. R. & Kan, Y. W. *Lancet* **i,** 3-5 (1986).
19. Nussinov, R. *J. molec. Biol.* **149,** 125-131 (1981).
20. Salser, W. *Cold Spring Harb. Symp. quant. Biol.* **42,** 985-1002 (1977).
21. Bird, A. P. *Nucleic Acids Res.* **8,** 1499-1504 (1980).
22. Coulondre, C., Miller, J. H., Farabaugh, P. J. & Gilbert, W. *Nature* **274,** 775-780 (1978).
23. Barker, D., Schafer, M. & White, R. *Cell* **36,** 131-138 (1984).

The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein

M. Z. Papiz[*†], L. Sawyer[*‡§], E. E. Eliopoulos[‖], A. C. T. North[‖], J. B. C. Findlay[¶], R. Sivaprasadarao[¶], T. A. Jones[*], M. E. Newcomer[*] & P. J. Kraulis[*]

[*] Department of Applied Chemical Sciences, Napier College, Edinburgh EH10 5DT, UK
[‡] Department of Biochemistry, University of Edinburgh, Edinburgh EH8 9XD, UK
[‖] Astbury Department of Biophysics, University of Leeds, Leeds LS2 9JT, UK
[¶] Department of Biochemistry, University of Leeds, Leeds LS2 9JT, UK
[*] Department of Molecular Biology, Box 590, BMC University of Uppsala, Uppsala, Sweden

Since its first isolation[1], bovine β-lactoglobulin (BLG) has been an enigma: although it is abundant in the whey fraction of milk, its function is still not clear. The results of the many physico-chemical studies on the protein need a structural interpretation. We report here the structure of the orthorhombic crystal form of cow BLG at pH 7.6, at a resolution of 2.8 Å. It has an unusual protein fold, composed of two slabs of antiparallel β-sheet, which shows a remarkable similarity to plasma retinol-binding protein. A possible binding site for retinol in BLG has been identified by model-building. This suggests a role for BLG in vitamin A transport and we have discovered specific receptors for the BLG–retinol complex in the intestine of neonate calves.

BLG is found in the milk of many animals including bovine species, sheep, deer, horse, dog, pig, dolphin[2] and possibly man. Usually the protein exists as a dimer and many species appear to possess genetic variants. In ruminants, the relative molecular weight of the subunit is ~18,000 corresponding to a polypeptide of ~162 residues. There are two disulphide bridges and, in the cow protein, a free cysteine in the sequence Cys-Gln-Cys- (119–121)[3], variously reported as 119[4], 121[5,6] or an equal mixture of the two[7]. The properties of the protein have been reviewed[4,8–10]. It is remarkably acid-stable, resisting denaturation at pH 2[11] and remaining intact until after passing through the stomach[12]. As the pH is raised, it undergoes several conformational changes observed by optical rotatory dispersion, perhaps the most interesting occurring at about pH 7.5[13]. BLG is known to bind a variety of hydrophobic molecules[14–16].

Of the abundant crystal forms of BLG[17,18], four salted-out forms have been studied by X-ray crystallography at low resolution[19]. A preliminary report has been given[20] of higher resolution work on one, lattice Y, crystallised at pH 7.6 We now discuss the similarity between BLG and plasma retinol-binding protein (RBP), which has suggested a function for BLG.

RBP is the protein that transports retinol from its storage in the liver. It is secreted into the plasma where it binds to transthyretin (formerly prealbumin). The transfer of retinol from plasma to retina is mediated by a specific surface receptor[21,22]. Human RBP contains 182 amino acids; the 2 Å resolution structure of the RBP-retinol complex has recently been solved[23] and refined[24].

We have determined the structure of BLG by standard crystallographic techniques (Fig. 1). We are confident that the strands are accurately positioned in our map, although one external loop and the N- and C-terminal residues remain poorly defined. Higher resolution data now being processed should reveal the remainder of the molecule. The molecule consists of anti-parallel β-sheet, formed by 9 strands wrapped round to form a flattened cone or calyx. The core of the molecule, an eight-stranded, antiparallel β-barrel, is an unusual structural motif which has been observed previously only in RBP[23], although catalase[25] has a barrel of similar topology, one connecting loop of which forms a separate domain. The strands A–I comprise residues 16–27, 39–44, 47–58, 62–76. 80–84, 89–97, 102–109, 115–124 and 145–150. Strands E,F,G,H,A and I are flanked by a 3-turn helix (130–140). The strands of the two sheets that form opposite sides of the calyx are orthogonal (ref. 26 and Fig. 2). Reverse turns occur at residues 44–47, 59–62, 78–81 and 84–88 but the turns at 98–102 and 111–115 are not true β-bends. Strand I is involved in the formation of the dimer by making anti-parallel interactions with the dyad-related strand. The interface between monomers also involves hydrophobic interactions between Ile 29 and Ile 147, and stacking of the imidazoles of symmetry-related His 146 residues. The free sulphydryl, at position 121, is buried at the sheet-helix interface. The disulphide bond at 106–119 is clearly resolved and shows

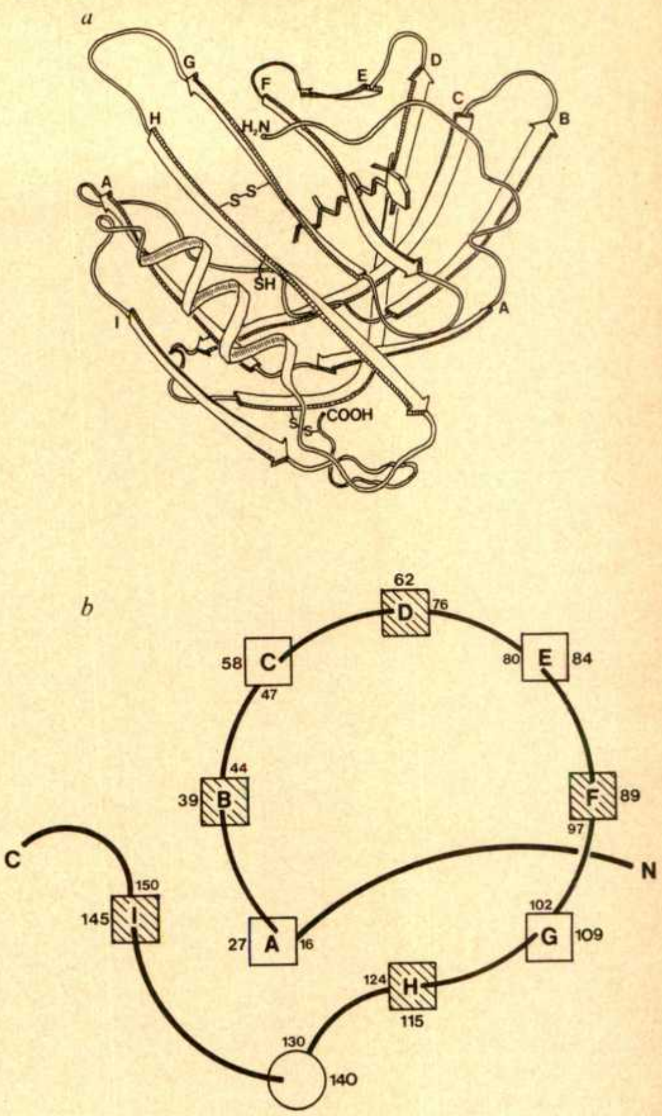


Fig. 1 *a*, Drawing of the β-lactoglobulin fold showing the position at which we propose that retinol binds. The arrows indicate the strands of β-sheet. The axis of the calyx runs approximately parallel to the extended isoprene chain with the far tip lying at the entrance from the solvent. *b*, The topology of the protein with the N-termini of the strands hatched to emphasize their antiparallel nature. A preliminary account of the preparation and crystallography of BLG is given in ref. 20.

---

† Present address: SERC Daresbury Laboratory, Warrington WA4 4AD, UK.
§ To whom correspondence should be addressed.

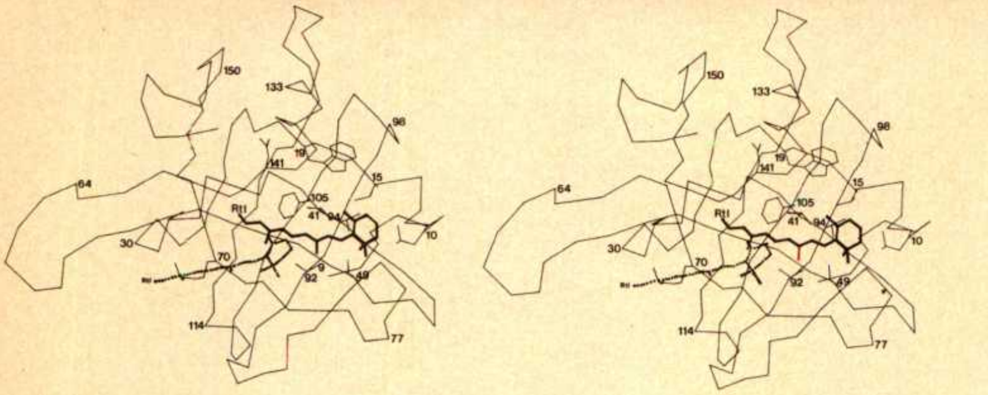


Fig. 2 Stereoscopic diagram of the superposition of the main chains of BLG (heavy) and RBP (light) viewed in such a way as to show the striking perpendicular packing of the β-sheets.

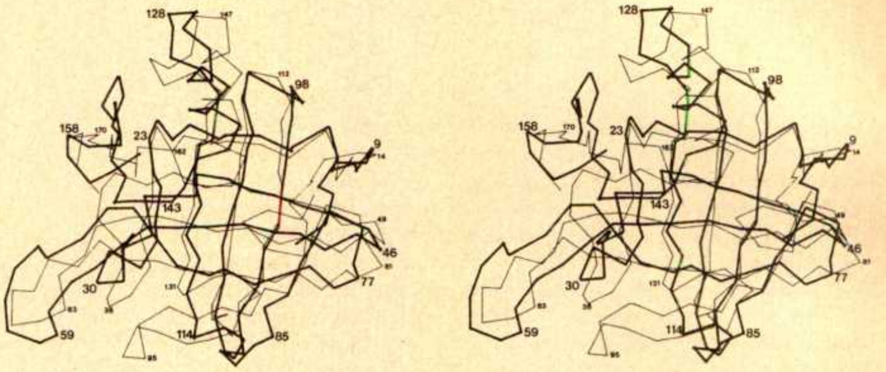


Fig. 3 Stereoscopic diagram of retinol (Rtl) model-built into β-lactoglobulin as described in the text. The side-chains of the BLG residues referred to are also shown as is the position of the retinol molecule (dotted) in RBP.

no evidence of equimolar amounts of cystine 106–119 and 106–121[7]. As the distance between the S_γ atoms of Cys 119 and Cys 121 is 10.5 Å, any interchange would be impossible without significant conformational flexibility. The other disulphide, 66–160, is not clear although it appears to be on the outer surface linking strand D to the C-terminal region.

BLG shows limited sequence homology to RBP, and human α_1- and rat α_2-microglobulins (refs 2, 27 and L. Rask and P. A. Peterson, personal communication). The sequence alignment of BLG and RBP, obtained with the program GAP[28] using a gap weight of 1 and a gap length of 0.1, shows 42 identities when 9 insertions or deletions are allowed. Only two pairs of regions have 4 consecutive identical residues (17–20 in BLG and 22–25 in RBP; 96–99 in BLG and 108–111 in RBP), and a third pair (79–83 in BLG and 83–87 in RBP) has 4 out of 5. However, the spatial similarity of the two structures is striking (Fig. 2) and when the backbone C_α coordinates were superimposed by the program HOMO[29,30], 129 out of the 162 residues were closely aligned, with an r.m.s. error of 2.76 Å. The greater length of the RBP chain is accommodated in larger surface loops, while the regular cores of the molecules are similar and in each case form a pocket open to the surface in which retinol binds in RBP.

Futterman and Heller[15] studied the binding of retinol to BLG and RBP by fluorescence spectroscopy and suggested that the retinol binding site in RBP was non-polar. In RBP the fluorescence remained constant with time whereas BLG showed a slow decline. The retinol was fully protected against oxidation by liver alchohol dehydrogenase in the complex with RBP but not in that with BLG. It has been suggested[31] that the BLG binding site makes the retinol more rigid than in free solution by encompassing the β-ionone ring and one or more adjacent carbon–carbon double bonds and that a tryptophan residue helps to form the binding site for the β-ionone moiety. Binding studies of retinylidene amine[32] suggest that both BLG and RBP protect the retinyl Schiff's base from hydrolysis and that in BLG there is a positively charged group near the carbon–nitrogen double bond. Interestingly, the dissociation constant, K_d, for the retinol-BLG complex[31] is 2×10^{-8} M whereas that for the RBP complex[33] is only 1.7×10^{-7} M.

These studies are in agreement with our structural models. In RBP[24], the retinol molecule is surrounded by the protein except for the very tip of the hydroxyl group. We modelled the retinol–BLG complex, first transforming the coordinates of retinol from RBP to the BLG coordinate system. Interactive graphics[34,35] gave a model in which the isoprene tail of retinol is fairly accessible to solvent because the loop at positions 83–85 in BLG is shorter than the corresponding one (at positions 91–98) in RBP. No tryptophan residue is close to the retinol although Trp 19, at the bottom of the calyx, lies at a distance of about 10 Å. Lys 70 may approach the retinol hydroxyl group. However, in BLG there appears to be an empty space further into the calyx and inaccessible to solvent, owing to changes from the residues that block this end of the cavity, in RBP. In particular, of 5 phenylalanine residues in RBP (at positions 15, 20, 45, 77 & 86) there are 3 (at positions 15, 20 and 45) that are replaced by smaller ones in BLG (Leu 10, Val 15 & Val 41); in addition, RBP Met 53 is changed to Thr 49 in BLG, RBP His 104 to BLG Val 92, RBP Ile 106 to BLG Val 94 and RBP Gln 117 to the larger Phe 105 in BLG. These changes allow the β-ionone ring of retinol to be fitted more deeply after minor conformational adjustments to some protein sidechains. The new binding site is partly formed by the sidechain of Trp 19. Most of the retinol is as well protected from the solvent in this conformation of BLG as in RBP, although the tip of the isoprene tail remains more exposed. It seems reasonable to imagine that some exchange is possible with the hydrophobic substrate binding pocket of ADH[36], which in the presence of NAD^+ could lead to oxidation. The retinol hydroxyl group is still within reach of the N_ζ of Lys 70. Energy minimization of the retinol in this deeper position reveal a better fit when the conjugated π-system is fully planar than when rotation is allowed about the 6-7 bond to permit closer approach to Trp 19. The fine structure in the UV spectrum of retinol when bound to BLG has been interpreted[37] as indicating a fully planar π-system. Thus, this model (Fig. 3) accounts for all the available chemical evidence, although details of the interactions between BLG and retinol will have to await the high resolution refined structure.

The fact that retinol binds to BLG with greater affinity than to RBP and our demonstration that it can be neatly accommodated within the BLG structure, prompted the speculation[20] that the biological function of the milk protein is involved with vitamin A transport. We monitored [^{3}H]-retinol bound to BLG using both gel filtration and a transthyretin-affinity resin, but found no detectable interaction between BLG and transthyretin. However, [^{125}I]-labelled BLG-retinol complex binds specifically to purified microvilli prepared from one-week-old calf intestine. Total binding of [^{125}I]-labelled-BLG (5 nM) was reduced by 70% in the presence of 5 μM unlabelled BLG. Specific binding was observed only in the lower segments of the small intestine and was not present in six-month old animals. The binding is dependent on both pH and ionic strength, and shows optima in the region of 6.8 and 0.1 M respectively. These observations are still preliminary, but they suggest that specific receptors for BLG which may be important in vitamin uptake are present in the small intestine of the neonate calf. The mechanism of uptake remains to be established. Neither RBP nor BLG inhibits the binding of the other to its specific receptor.

We thank Professor Sir David Phillips for drawing our attention to the similarity between β-lactoglobulin and RBP, Dr C. Holt for help with protein purification, Drs H. Muirhead and H. C. Watson, Bristol University, for use of X-ray photographic facilities, the Physics Department, Edinburgh University for use of their diffractometer and computing facilities and the SERC for financial support. E.E.E. thanks the University of Leeds for a studentship.

Received 17 February; accepted 30 September 1986.

1. Palmer, A. H. *J. biol. Chem.* **104,** 359-372 (1934).
2. Pervaiz, S. & Brew, K. *Science* **228,** 335-337 (1985).
3. Grosclaude, F., Mahe, M. F., Mercier, J. C., Bonnmaire, J. & Teissier, J. H. *Ann. Genet. Sel. anim.* **8,** 461-479 (1976).
4. McKenzie, H. A. in *Milk Proteins: Chemistry and Molecular Biology* Vol. 2 (ed. McKenzie, H. A.) 257-330 (Academic, New York, 1971).
5. Martial, J., Preaux, G. & Lontie, R. *Arch. int. Physiol. Biochim.* **79,** 842-843 (1971).
6. Mainferme, F., Preaux, G. & Lontie, R. *Arch. int. Physiol. Biochim.* **79,** 840-841 (1971).
7. McKenzie, H. A., Ralton, G. B. & Shaw, D. C. *Biochemistry* **11,** 4539-4547 (1972).
8. Lyster, R. L. J. *J. Diary Res.* **39,** 279-318 (1972).
9. Jenness, R. in *Lactation: a Comprehensive Treatise,* Vol. 3, (eds Larson, B. L. & Smith, V. R.) 3-107 (Academic, New York, 1974).
10. Jenness, R. *J. Diary Res.* **46,** 197-210 (1979).
11. Aschaffenburg, R. & Drewry, J. *Biochem. J.* **65,** 273-277 (1957).
12. Miranda, G. & Pelissier, J. P. *J. Dairy Res.* **50,** 27-36 (1983).
13. Tanford, C., Bunville, L. G. & Nozaki, Y. *J. Am. chem. Soc.* **81,** 4032-4036 (1959).
14. Lovrien, R. & Anderson, W. F. *Arch. Biochem. Biophys.* **131,** 139-144 (1969).
15. Futterman, S. & Heller, J. *J. biol. Chem.* **247,** 5168-5172 (1972).
16. Stone, W. L. & Wishnia, A. *Bioinorg. Chem.* **8,** 517-529 (1979).
17. Aschaffenburg, R., Green, D. W. & Simmons, R. M. *J. molec. Biol.* **13,** 194-210 (1965).
18. Bolognesi, M., Liberatori, J., Oberti, R. & Ungaretti, L. *J. molec. Biol.* **131,** 411-413 (1979).
19. Green, D. W. *et al. J. molec. Biol.* **131,** 375-397 (1979).
20. Sawyer, L., Papiz, M. Z., North, A. C. T. & Eliopoulos, E. E. *Biochem. Soc. Trans.* **13,** 265-266 (1985).
21. Rask, L. *et al. Scand. J. clin. Lab. Invest.* **40,** Suppl. 154, 45-61 (1980).
22. Goodman, D. S. *The Retinoids* Vol. 2, 41-48 (1984).
23. Newcomer, M. E. *et al. EMBO J.* **3,** 1451-1454 (1984).
24. Jones, T. A. & Newcomer, M. E. *J. molec. Biol.* (submitted).
25. Murthy, M. R. N., Reid, T. J., Sicgnano, A., Tanaka, N. & Rossmann, M. G. *J. molec. Biol.* **152,** 465-499 (1981).
26. Chothia, C. & Janin, J. *Biochemistry* **21,** 3955-3965 (1982).
27. Godovac-Zimmermann, J., Conti, A., Liberatori, J. & Braunitzer, G. *Hoppe-Seyler Z. Phys. Chem.* **366,** 431-434 (1985).
28. Devereux, J., Haeberli, P. & Smithies, O. *Nucleic Acid Res.* **12,** 387-395 (1984).
29. Rao, S. T. & Rossmann, M. G. *J. molec. Biol.* **76,** 241-256 (1973).
30. Rossmann, M. G. & Argos, P. *J. biol. Chem.* **250,** 7525-7537 (1975).
31. Fugate, F. D. & Song, P.-S. *Biochim. biophys. Acta* **652,** 28-42 (1980).
32. Horwitz, J. & Heller, J. *J. biol. Chem.* **249,** 4712-4719 (1974).
33. Cogan, U., Kopelman, M., Mokady, S. & Shinitzky, M. *Eur. J. Biochem.* **65,** 71-78 (1976).
34. Cambillau, C., Horjales, E. & Jones, T. A. *J. mol. Graph.* **2,** 53-54 (1984).
35. Jones, T. A. *J. appl. Cryst.* **11,** 268-272 (1978).
36. Eklund, H. *et al. J. molec. Biol.* **146,** 561-587 (1981).
37. Hemley, R., Kohler, B. E. & Siviski, P. *Biophys. J.* **28,** 447-455 (1979).

DNA conformation is determined by economics in the hydration of phosphate groups

Wolfram Saenger*, William N. Hunter† & Olga Kennard†

* Institut für Kristallographie Freie Universität Berlin, Takustr. 6, D-1000 Berlin 33, FRG
† University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK

Mixed sequence DNA can exist in two right-handed and one left-handed double helical conformations—A, B and Z[1–3]. Under conditions of high water activity the B conformation prevails. If the water activity is reduced on addition of salt or organic solvents, transformation occurs to A-DNA or, in DNAs with alternating purine-pyrimidine sequences, to the left-handed Z-DNA. In crystal structure analyses of oligonucleotides, the free oxygen atoms of adjacent phosphate groups along the polynucleotide chain in B-DNA are found at least 6.6 Å apart and individually hydrated[4] whereas they are as close as 5.3 Å in A-DNA and 4.4 Å in Z-DNA, and bridged by water molecules[5–7]. We suggest that this more economical hydration in A- and Z-DNA compared with B-DNA is the underlying cause of B → A and B → Z transitions.

To learn about possible driving forces for conformational transitions of DNA, we have investigated in detail the hydration of crystalline DNA fragments in the A, B and Z forms. The bases and sugars are equally well hydrated in all three forms. Ordered patterns of hydration have been observed in several oligonucleotide crystal structures, including the spines of hydration in the minor groove of the Z-hexamers with alternating G-C sequence[6–9], in the centre AATT part of the B-DNA dodecamer[4,10–13], and the fused pentagons within the TATA region of the A-DNA octamer[5]. These ordered structures are base-specific and are important elements in stabilizing regions of DNA with particular base sequences, and in influencing

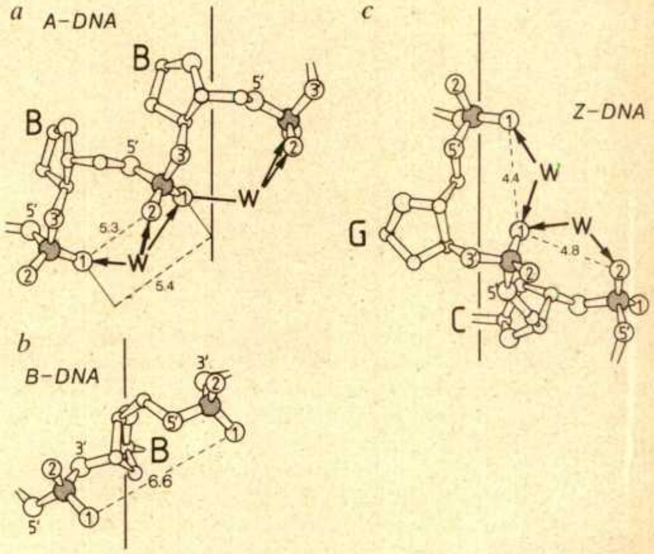


Fig. 1 Comparison of deoxyribose-3′, 5′-diphosphate fragments in *a*, A-, *b*, B- and *c*, Z-DNA. Redrawn from corresponding figures in ref. 1. Vertical lines represent helical axes; phosphorus atoms indicated by shading, phosphate oxygens by numbering. Shortest distances between free phosphate oxygen atoms are given in Å units (see Table 1), W represents water molecules that bridge free phosphate oxygens in A- and Z-DNA. Capital letters B in A- and B-DNA and G, C in Z-DNA indicate C(1′) sugar atoms where bases are attached.

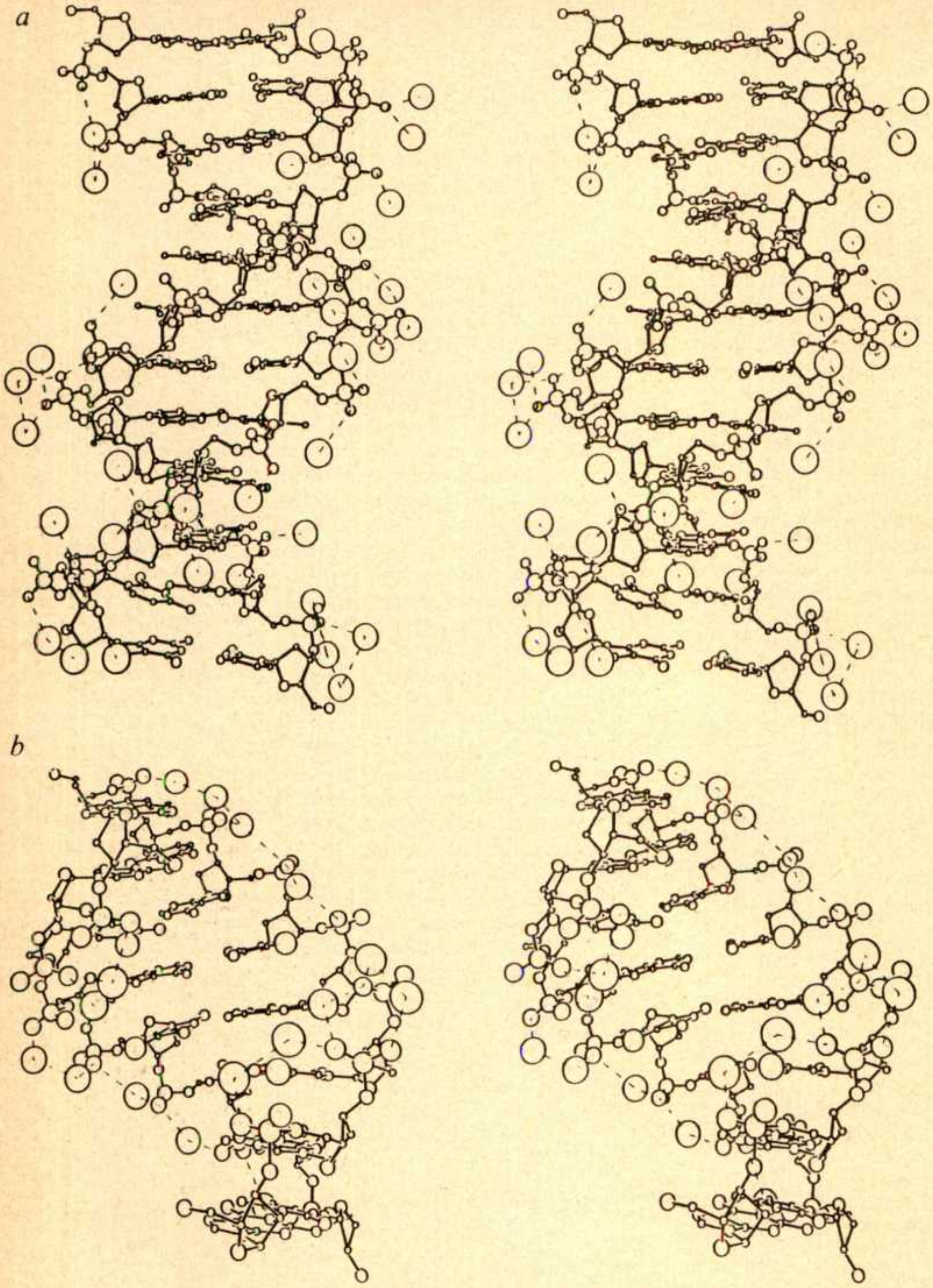


Fig. 2 *a*, *b*, DNA stereoviews. *a*, Stereoview showing the hydration of phosphate groups in the B-DNA dodecamer d(CGCGAATTAGCG) (ref. 10). Atoms are shown in order of size: water oxygen, P, O, N, C. Dashed lines, hydrogen bonds. Note the clustering of solvent molecules around phosphate groups and the lack of continuous water bridges. *b*, Hydration of phosphate groups in the A-DNA fragment d(GG$^{\mathrm{Br}}$UA$^{\mathrm{Br}}$UACC) (ref. 5). Of the 12 possible water bridges between free oxygen atoms of adjacent phosphate groups, 7 are actually formed within the cutoff limit of 3.5 Å for O–O separation in water. In one strand of the duplex, the water bridges are nearly continuous. The stereoview is into the major groove.

conformational transitions of DNA with non-random sequence, such as the B → A transitions of DNA fibres with high A+T content[14], and the B → Z transition[15] of poly [d(G-C)]. However, it is not yet possible to extrapolate from these sequence-dependent ordered patterns to the more general mechanisms of conformational transitions in mixed-sequence DNA.

The most relevant difference between A-, B- and Z-DNA lies in the conformation of the sugar ring, which is related to the distance between the phosphate groups at the 3′ and 5′ ends of each nucleotide, and to the displacement of base pairs relative to the helical axis. The sugar pucker in B-DNA is 'soft' and extends beyond the O(4′)-*endo* to C(1′)-*exo* to C(2′)-*endo* range or, in terms of pseudorotation phase angle[16], from east to south, with the majority of sugars in C(1′)-*exo* to C(2′)-*endo*. These sugar puckers give rise to a wide phosphate–phosphate separation of 6.6 Å on average (Table 1 and Fig. 1). In contrast, in A-DNA the sugar conformation is much more narrowly confined to C(3′)-*endo*, and the distance between the phosphate groups is reduced to an average of about 6 Å. In Z-DNA, the sugar pucker alternates, with C(2′)-*endo* for the pyrimidine nucleotides in *anti* conformation, and with O(4′)-*endo* to C(1′)-*exo* for the *syn* purine nucleotides. The resulting interphosphate distances differ considerably for the sequences listed in Table 1, but are comparable to A-DNA in the idealized Z-DNA, 5.9 Å for the pCp step and 6.0 Å for the pGp step.

Associated with the differences in phosphate–phosphate separation in A-, B-, Z-DNA are differences in distances between the successive free phosphate oxygens O(1) and O(2), that are further influenced by right- and left-handed helical twist, as shown in Table 1 and Fig. 1. On average, all interphosphate O–O distances exceed 6.6 Å in B-DNA. In A-DNA (omitting the very short sequence containing an iodinated cytidine residue d($^{\mathrm{I}}$CCGG)), they are considerably smaller; 5.4–5.7 Å for O(1)–

Table 1 Average intra-strand separations (Å) between free oxygen atoms O(1), O(2) in adjacent phosphate groups in A-, B-, Z-DNA

Type	Sequence	P-P	O(1)-O(1)	O(2)-O(2)	3′O(2)-O(1)5′	3′O(1)-O(2)5′	Resolution (Å)	Reference
		Intra-strand atomic distances (Å)						
A	d(GGGGCCCC)	6.0-6.6 6.2(0.8)	5.0-6.2* 5.7(0.9)	5.9-7.3 6.7(0.7)	7.0-7.9 7.6(0.8)	4.8-6.3* 5.5(0.8)	2.5	24
A	d(GG^{Br}UA^{Br}UACC)	5.6-6.3 6.0(0.7)	4.9-6.0* 5.5(0.7)	6.2-7.1 6.6(0.7)	7.0-7.9 7.4(0.8)	5.0-6.2* 5.5(0.8)	1.7	5
A	d(GGGGTCCC)	5.7-6.2 5.9(0.3)	5.0-6.0* 5.4(0.3)	6.3-6.7 6.5(0.1)	7.1-7.7 7.3(0.2)	4.9-5.8* 5.3(0.3)	2.2	25
A	d(GGGGCTCC)	5.5-6.6 5.8(0.5)	5.1-6.4* 5.6(0.5)	5.9-7.0 6.0(0.6)	6.8-7.9 7.5(0.4)	4.5-5.9* 5.4(0.5)	2.2	26
A	d(^{I}CCGG)	5.5-6.6 6.3(0.6)	5.9-7.3 6.6(0.6)	4.9-6.5 6.1(0.7)	7.0-8.1 7.6(0.6)	4.4-6.4 5.9(0.8)	2.2	27
B	d(CGCGAATTCGCG)	6.3-7.1 6.7(0.8)	5.9-7.5 6.8(0.8)	6.1-8.4 7.4(0.9)	6.7-8.9 8.1(1.0)	5.7-7.7 6.7(1.0)	2.3	11
B	d(CGCGAATTAGCG)	6.3-7.1 6.6(0.9)	5.8-7.8 6.7(0.9)	5.8-8.0 7.3(1.0)	7.1-9.0 8.0(1.0)	6.0-7.6 6.7(1.0)	2.5	10
B	d(CGCGAATTTGCG)	6.0-7.2 6.6(0.3)	6.3-8.3 6.7(0.5)	6.1-8.7 6.9(0.6)	7.0-9.1 8.1(0.6)	5.8-7.3 6.6(0.6)	2.5	12
B	d(CGCAAATTCGCG)	5.9-7.1 6.6(0.5)	4.8-8.0 6.6(0.7)	5.9-9.0 7.4(0.7)	6.2-9.0 7.6(0.8)	5.8-8.3 6.7(0.6)	2.5	13
Z	d(^{Br}UGCGCG)							
	step pCp	6.1-6.2 6.1(0.1)	7.0-7.3 7.2(0.1)	5.8-6.5 6.2(0.3)	5.0-5.2* 5.1(0.1)	8.1-8.2 8.1(0.1)		
	step pGp	7.0-7.2 7.1(0.1)	5.3-5.8* 5.6(0.3)	8.8-9.6 9.2(0.4)	7.0-8.2 7.6(0.6)	7.2-7.6 7.5(0.4)	2.2	28
Z	d(TGCGCG)							
	step pCp	5.8-6.0 5.9(0.1)	6.5-7.0 6.8(0.3)	5.7-5.9 5.8(0.1)	4.4-5.0* 4.7(0.3)	7.7-7.8 7.8(0.1)		
	step pGp	6.3-6.9 6.5(0.3)	4.6-6.2* 5.2(0.8)	8.5-9.0 8.7(0.3)	6.7-7.0 6.8(0.2)	6.7-7.4 7.0(0.4)	1.5	12
Z	d(^{Br}CG^{Br}CG^{Br}CG)							
	step pCp	5.5-5.9 5.8(0.2)	6.1-6.7 6.5(0.3)	4.7-6.0 5.5(0.6)	3.9-4.6* 4.3(0.3)	7.1-7.9 7.6(0.4)		
	step pGp	6.1-6.5 6.3(0.2)	4.7-6.0* 5.2(0.6)	8.1-8.7 8.5(0.3)	6.7-7.3 7.1(0.2)	6.4-7.0 6.8(0.3)	1.5	7
Z	d(CGCGCG) ideal							
	step pCp	5.9	7.0	5.5	4.8*	7.6		15
	step pGp	6.0	4.4*	8.2	6.4	6.3		

In A- and B-DNA, O(1) is that atom which is closer to the major groove side. Numbers give the spread of distances and the geometrical mean with standard deviation in parentheses. In the DNAs with base pair mismatches, the backbone geometry is virtually unchanged compared with all-Watson-Crick base pairs.

* Oxygen-oxygen contacts which are so close that they are most frequently (and almost systematically) bridged by water molecules. See also Fig. 1.

Table 2 Statistics of water binding to free phosphate oxygen atoms O(1), O(2) in different DNA fragments

Sequence*	I	II	III	IV	V	VI	VII
Helix type	A	A	B	B	B	Z	Z
Number of water molecules:							
Bound to O(1), O(2)	40	34	22	35	32	37	35
Forming bidentate intra-phosphate bridges	1	0	1	2	2	1	1
Forming monodentate interactions with O(1), O(2)	31	30	21	33	30	34	31
In one-water inter-phosphate bridges	7	4	0	0	0	5	3
In two-water inter-phosphate bridges	4	1	1	2	1	2	2
Inter-phosphate gaps bridged by one water (%)	58	33	0	0	0	63	38

The cutoff limit for O-O distances to water is 3.5 Å.

* I is d(GG^{Br}UA^{Br}UACC), ref. 5; II is d(GGGGCCCC), ref. 24; III is native d(CGCGAATTCGCG) ref. 11; IV is d(CGCGAATT^{Br}CGCG), MPD 7, ref. 4; V is d(CGCGAATTAGCG), ref. 10; VI is d(^{Br}CG^{Br}CG^{Br}CG, ref. 7; VII is d(^{Br}UGCGCG), ref. 28.

O(1), and 5.3 Å for 3′-O(1)-O(2)-5′, compared to the other two distances O(2)-O(2) and 3′-O(2)-O(1)-5′. For Z-DNA in the idealized form, O(1)-O(1) distance for pGp (4.4 Å) and 3′-O(2)-O(1)-5′ distance for pCp (4.8 Å) are short; O(2)-O(2) for pCp (5.5 Å) is intermediate; and all other O-O distances exceed 6.4 Å.

When the hydration schemes of the phosphate groups in the three different DNA conformations are compared, it is evident that in B-DNA the average oxygen-oxygen separation of about 6.6-8.1 Å is so large that each phosphate group can be surrounded by its own sphere of water molecules, and displays an independent hydration[4]. With a cutoff of 3.5 Å for O-O$_{water}$ distances, one or two water molecules cannot bridge adjacent phosphates for the sequences given in Table 2. In contrast, in A- and Z-DNA, the oxygen atoms of adjacent phosphate groups are so close together (Table 1) that they can be bridged by one or two water molecules[5-7], as shown in Figs 1, 2 and in Table 2. The single water molecule bridges are always oriented in the same direction; in the 3′ to 5′ direction in A-DNA, the O(1) of the 3′-phosphate being linked by a water molecule to either the O(1) or O(2) oxygens of the adjacent 5′-phosphate, as these O-O distances are the shortest. In crystal structures the former is preferred because only in this case can the water molecule also interact with the functional groups of the bases in the major groove. For Z-DNA, the reverse direction of the bridges is found in the pCp step, O(2)-O(1) (Fig. 1). This directionality of water bridges is a result of the right- and left-handed twists of the respective DNA conformers, which determine the separations between phosphate oxygens.

B-DNA occurs under conditions of high water activity, and

A- and Z-DNA at low water activity; it appears that the hydration of A- and Z-DNA, where adjacent phosphate groups along the polyphosphate backbone are bridged by water molecules, is more economical and efficient than that of B-DNA, where each phosphate group is individually hydrated. As indicated by gravimetric[17] and infrared spectroscopic[18] studies and by crystal structure analyses of hydrated nucleotides[19], the binding of water molecules to the charged phosphate oxygen atoms is energetically more favoured than binding to base and sugar atoms. Therefore, the phosphate hydration will have a substantial influence on the structure of DNA. If the water activity is high, the number of water molecules available for hydration of DNA is high, and the phosphate groups will be fully and individually hydrated. If salt or an organic polar solvent is added, water molecules are withdrawn from the DNA and hydration consequently becomes more economical. Economy is achieved by water molecules bridging adjacent phosphate groups, following a change in conformation of the sugar residues that results in the phosphate groups moving closer together. This transition between conformational states is cooperative, as is the formation of the bridges, which run along the entire sugar-phosphate backbone in A-DNA[5]. A comparable situation is seen in Z-DNA[6,7] and we assume that this transition has the same basis as the B → A transition. Additionally, in Z-DNA the metal ions appear to play a significant role as in the crystal structures magnesium hexahydrate, cobalt hexammine, and sodium ions were found tightly coordinated with N(7) and O(6) of guanosine[7,9,20].

The concept of hydration economy finds further support if we consider the water content per base pair in the DNA double helices. It has been suggested, on the basis of gravimetric studies[21] and theoretical considerations[22], that an A · T base pair binds one more water molecule than a G · C base pair. This was confirmed in the crystal structure of the B-DNA dodecamer where water molecules are systematically trapped between the thymine methyl groups and one of the free oxygens of the 5′ phosphates[23]. Such an interaction can only occur if the nucleotide adopts a sugar conformation typical of B-DNA because then the base and the 5′-phosphate are farther apart than when the sugar is in the C(3′)-*endo* conformation found in A-DNA. Therefore, DNA at high water activity adopting the B-type conformation is able to bind one additional water molecule at each thymidine, which confers additional stabilization. If the water activity is lowered, this (weakly) bound water molecule will be removed, and the transition from the B-type sugar puckering mode to the C(3′)-*endo* form can take place. This again has to occur under cooperative control because otherwise the DNA would have both random B-DNA type and C(3′)-*endo* puckers which would be unfavourable for the double helical structure.

Hence, it appears that the conformational transitions of mixed sequence DNA from B to A or B to Z are associated with the hydration of the phosphate groups which is less economical in B-DNA than in the A and Z forms. It is clear that other factors such as the spine of hydration in the minor groove of B-DNA and Z-DNA and the ordered water pentagons in the major groove of A-DNA also play a role, particularly in stabilizing specific base sequences.

O.K. is a member of the scientific staff of the Medical Research Council and the work was supported by an MRC programme grant. W.S. acknowledges a short-term EMBO fellowship to visit the University Chemical Laboratory, and support by Sonderforschungsbereich 9. We thank Dr E. Westhof for providing coordinates for Z-DNA and Dr R. E. Dickerson for constructive criticism of this paper.

Received 7 April; accepted 12 September 1986.

1. Saenger, W. in *Principles of Nucleic Acid Structure*, 220-297 (Springer, New York, 1983).
2. Dickerson, R. E. *et al. Science* **216**, 475-485 (1982).
3. Jurnak, F. A. & McPherson, A. (eds) *Biological Macromolecules and Assemblies* Vol. 2: *Nucleic Acids & Interactive Proteins* (Wiley, New York, 1985).
4. Kopka, M. L., Fratini, A. V., Drew, H. R. & Dickerson, R. E. *J. molec. Biol.* **163**, 129-146 (1983).
5. Kennard, O. *et al. J. biomolec. struct. Dyn.* **3**, 205-225 (1986).
6. Westhof, E., Prangé, Th., Chevrier, B. & Moras, D. *Biochimie* **67**, 811-817 (1985).
7. Chevrier, B. *et al. J. molec. Biol.* **188**, 707-719 (1986).
8. Wang, A. H.-J., Hakoshima, T., van der Marel, G., van Boom, J. H. & Rich, A. *Cell* **37**, 321-331 (1984).
9. Quigley, G. J. in *Molecular Structure and Biological Activity* (eds Griffin, J. & Duax, W. L.) 317-332 (Elsevier, New York, 1981).
10. Brown, T., Hunter, W. N., Kneale, G. & Kennard, O. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2402-2406 (1986).
11. Dickerson, R. E. & Drew, H. R. *J. molec. Biol.* **149**, 761-786 (1981).
12. Kennard, O. *J. molec. struct. Dyn.* **3**, 205-225 (1985).
13. Hunter, W. N., Brown, T., Anand, N. N. & Kennard, O. *Nature* **320**, 552-555 (1986).
14. Arnott, S. & Selsing, E. *J. molec. Biol.* **88**, 509-521 (1974).
15. Wang, A. H.-J. *et al. Science* **211**, 171-176 (1981).
16. Altona, C. & Sundaralingam, M. *J. Am. chem. Soc.* **94**, 8205-8212 (1972).
17. Falk, M., Hartmann, K. A. Jr & Lord, R. C. *J. Am. chem. Soc.* **84**, 3843-3846 (1962).
18. Falk, M., Hartmann, K. A. Jr & Lord, R. C. *J. Am. chem. Soc.* **85**, 387-391 (1963).
19. Jeffrey, G. A., Maluszynska, M. & Mitra, J. *Int. J. biol. Macromol.* **7**, 336-347 (1985).
20. Gessner, R. V. *et al. Biochemistry* **24**, 237-240 (1985).
21. Tunis, M.-J. B. & Hearst, J. E. *Biopolymers* **6**, 1345-1353 (1968).
22. Goldblum, A., Perahia, D. & Pullman, A. *FEBS Lett.* **91**, 213-315 (1978).
23. Drew, H. R. & Dickerson, R. E. *J. molec. Biol.* **151**, 535-556 (1981).
24. McCall, M., Brown, T. & Kennard, O. *J. molec. Biol.* **183**, 385-396 (1985).
25. Kneale, G., Brown, T., Kennard, O. & Rabinovich, D. *J. molec. Biol.* **186**, 805-814 (1985).
26. Hunter, W. N., Kneale, G., Brown, T., Rabinovich, D. & Kennard, O. *J. molec. Biol.* (in the press).
27. Conner, B. N., Yoon, C., Dickerson, J. L. & Dickerson, R. E. *J. molec. Biol.* **174**, 663-695 (1984).
28. Brown, T., Kneale, G., Hunter, W. N. & Kennard, O., *Nucleic Acids Res.*, **14**, 1801-1809 (1986).

A protein on *Plasmodium falciparum*-infected erythrocytes functions as a transferrin receptor

Mario H. Rodriguez* & Michele Jungery†

Department of Tropical Public Health,
Harvard School of Public Health,
665 Huntington Avenue, Boston, Massachusetts 02115, USA

Several observations suggest that iron is essential for the development of malaria parasites[1-6] but there is evidence that the parasites in erythrocytes do not obtain iron from haemoglobin. The total haemin level in parasitized erythrocytes does not vary during parasite development[7], indicating that the iron-containing moiety of haemoglobin is not detectably metabolized. Although parasite proteases can degrade the protein part of haemoglobin in red cells[8,9], no parasite enzymes that degrade haemin have been identified. In mammalian cells, haemin is degraded to carbon monoxide and bilirubin by the enzyme haeme oxygenase[10]. This enzyme has not been found in malaria parasites[11]. In fact haemin has been found to be toxic to parasite carbohydrate metabolism[12]. Thus, iron apparently cannot be liberated from haemin and instead is sequestered in infected red cells as haemozoin, the characteristic pigment associated with malarial infection[13,14]. If iron bound to transferrin is the source of ferric ions for malaria parasites within mature erythrocytes, then the parasite must synthesize its own transferrin receptor and localize it on the surface of the infected cell, because the receptors for transferrin are lost during erythrocyte maturation[5,16]. Our results here suggest that *Plasmodium falciparum* synthesizes its own transferrin receptors enabling it to take up iron from transferrin by receptor-mediated endocytosis.

To determine whether human ferrotransferrin binds to parasitized cells, we conjugated this ferrotransferrin with fluorescein isothiocyanate (FITC) and observed the movement of the ligand from the cell surface to the parasitophorous vacuole inside erythrocytes for 30 minutes (Fig. 1). The transferrin was initially found in small vesicles in the cytoplasm of infected cells but within 30 min these fluorescent vesicles concentrated around the

---

* Present address: Centro de Investigaciones de Paludismo, Tapachula, Chiapas, Mexico 30700.
† To whom correspondence should be addressed.

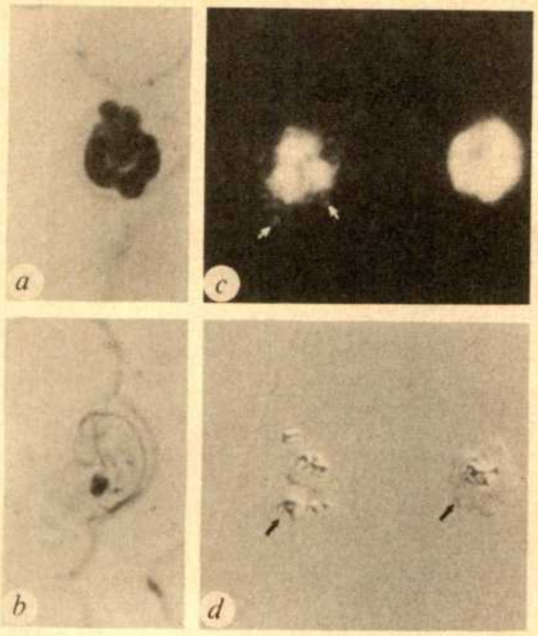


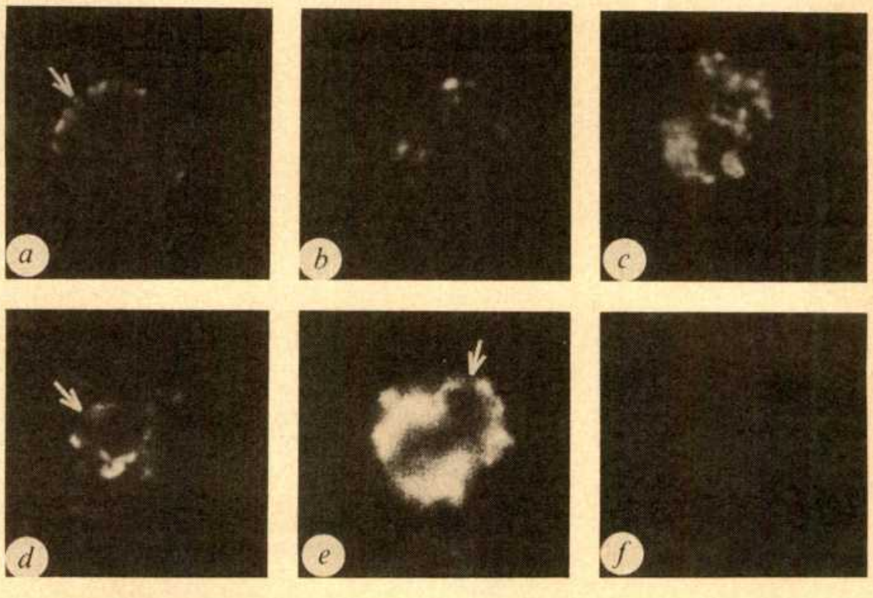
Fig. 1 Uptake of FITC-labelled ferrotransferrin by malaria parasites. Erythrocytes infected with *P. falciparum* (Palo Alto, Uganda strain) young trophozoites were obtained from cultures[26] by sedimentation on plasmagel[27]. The infected cells were washed twice with RPMI 1640 (GIBCO) and 25 μl of these cells were cultured in 100 μl of complete culture medium (10% human serum, 0.4% glucose, 0.2% $NaHCO_3$, 0.3% TES (*N*-Tris[hydroxymethyl]methyl-2-aminoethanolsulphonic acid), 0.004% gentamycin in RPMI 1640 medium) with ~1 mg ml^{-1} of FITC-ferrotransferrin. Transferrin (Calbiochem) was saturated with iron[28] and coupled to FITC (isomer I-Sigma) as described[29] with a final ratio FITC/protein of 2:1. Every 15 min cells were removed from culture, washed three times with cold RPMI and examined in a Zeiss ICM microscope with an epifluorescence condenser. Fluorescent transferrin was observed initially in small vesicles in the cytoplasm of the infected cells (*a*, arrow *b* and *c*). After 30 min at 37 °C fluorescent material accumulated around the parasitophorous vacuole surrounding the parasite (*d* and *e*, arrows). Fluorescence was not evident on uninfected cells on the same slide. The specificity of FITC-transferrin uptake was further tested by showing competititve inhibition of uptake in the presence of an excess of unlabelled transferrin (a 1:100 dilution of the labelled transferrin): only a pale surface-fluorescence was observed (*f*).

parasitophorous vacuole surrounding the parasite in all the parasitized cells (Fig. 1*d*, *e*, arrows). Such fluorescence in parasitized cells was observable over a wide range of FITC-transferrin concentrations. When FITC-ferrotransferrin was diluted with an excess of unlabelled ferrotransferrin, little fluorescence was observed on the infected cells (Fig. 1*f*) indicating selective competition for this receptor. K562 cells, which carry transferrin receptors, were used to show the specificity of binding of FITC-transferrin. FITC-transferrin could be detected at concentrations from 1 mg ml^{-1} to less than 1 μg ml^{-1} except in the presence of unlabelled transferrin. Finally, iron-free apotransferrin, which has no affinity for transferrin receptor at neutral *p*H, did not bind to any of these cells.

We confirmed the presence of transferrin in vesicles and in or around the parasites in similar experiments using either peroxidase (Fig. 2*a*, *b*) or rhodamine (Fig. 2*c*) immunostaining of fixed parasite-infected cells treated with an anti-transferrin rabbit antibody. Most of the fluorescence was in or around the parasites or in vesicles (Fig. 2*c*, arrows) and not at the surface. These results agree with observations in mammalian cells that about 75% of the transferrin bound to receptors is found intracellularly[17]. Studies at the ultrastructural level to determine the precise location of the ligand and the receptor are in progress (D. Titus, M. Neutra, M. Jungery).

We used affinity chromatography and binding assays to identify the receptor(s) for transferrin in malaria-infected cells. Two different types of affinity chromatography were used to isolate the transferrin receptor. First, a ^{35}S-labelled parasite extract was passed through a column of anti-transferrin immunoglobulin coupled to Sepharose 4B. The SDS-polyacrylamide gel electrophoresis(PAGE) profile of parasite proteins from this column reveals a prominent band with relative molecular mass (M_r) of approximately 93,000 (93K), and a series of minor bands (Fig. 3, lane 2). In the second isolation procedure, ferrotransferrin was coupled to Sepharose 4B and washed in a way that releases the receptor from the transferrin/receptor complex in mammalian systems[18]. After incubation with the labelled parasite extract, the eluates contained few counts, indicating that most of the bound parasite material was not released. The major parasite protein that bound to the transferrin affinity column was again the 93K protein (Fig. 3, lane 4). The 93K protein did not bind to control columns coupled with either bovine serum albumin or ethanolamine. Therefore the most likely candidate for the parasite-derived transferrin receptor is the 93K transferrin-binding protein.

Fig. 2 *a*, *b*, Immunoperoxidase detection of transferrin on *P. falciparum*-infected erythrocytes; *c*, *d*, rhodamine-conjugated immunostains of *P. falciparum* infected cells. *P. falciparum* was cultured in complete medium containing 1 mg ml^{-1} ferrotransferrin. Infected cells were washed three times with RPMI and thin smears prepared. After fixation with 10% methanol in acetone, the slides were washed with 0.5% H_2O_2 to inactivate endogenous peroxidase[30], and treated with rabbit anti-transferrin serum and a peroxidase-labelled goat anti-rabbit IgG (Cappel). After development with aminobenzidine-H_2O_2, a red-brown stain was observed in infected erythrocytes around the parasitophorous vacuole (*a*). Specific anti-transferrin stain was seen on all infected cells. No stain was seen on non-infected cells or on infected cells treated with non-immune rabbit serum (*b*). *c*, *d*, Parasites were fixed and incubated with anti-transferrin antibody as in *a*. Slide 5 was washed in PBS with gelatin, then rhodamine-conjugated goat anti-rabbit serum was added. *c*, Anti-transferrin binding to vesicles (arrow) and accumulating around the parasite. *d*, A light-field (Nomarski optics) photograph of the same cells as in *c*. Only faint background fluorescence can be observed on normal erythrocytes seen in *c*.



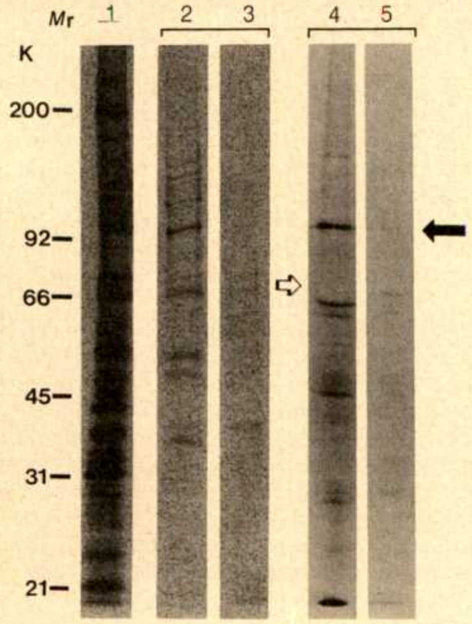


Fig. 3 Affinity chromatography of ^{35}S-methionine-labelled *P. falciparum* protein preparations on anti-transferrin-antibody bound to Sepharose or Fe-transferrin-Sepharose columns. *P. falciparum* trophozoites were metabolically labelled with ^{35}S-methionine for four hours. The parasites were harvested, solubilized in 0.15 M NaCl, 0.1 M PO$_4$, *p*H 7.4 in phosphate-buffered saline (PBS) with 100,000 U ml^{-1} aprotinin, 1 mM iodacetic acid, 5 mM TPCK, 5 mM TLCK and 2.0% Triton X-100. The preparation was incubated either with affinity purified rabbit anti-transferrin-antibody bound to Sepharose 4B or with Fe-transferrin-bound Sepharose 4B beads. Control columns included Sepharose 4B beads coupled with purified rabbit Ig (lane 3) or blocked with ethanolamine or bovine serum albumin (lane 5). Open arrow, location of the transferrin that was removed from the transferrin affinity Sepharose columns during boiling of the beads in SDS sample buffer.

Methods. The affinity column beads were incubated at room temperature for 3 h and then washed. The antitransferrin-Sepharose beads were washed with phosphate-buffered saline (PBS), 0.25% Triton X-100 until no counts were obtained from the washings. The Fe-transferrin beads were washed first with PBS, 0.5% Triton X-100, followed by 0.15 M Na acetate, 0.5% Triton X-100, *p*H 5, then 0.1 M Na citrate, 0.5% Triton X-100, 50 μg ml^{-1} Desferrioxamine, *p*H 5, and finally by 0.1 M NH$_4$HCO$_3$, 0.5% Triton X-200, 50 μg ml^{-1} Desferrioxamine, *p*H 7.8. Beads from both columns were then boiled for 5 min in sample buffer, (0.1 M Tris-HCl, *p*H 7.5, 2% SDS, 10% mercaptoethanol, 2 mM ethylenediamine-tetraacetic acid (EDTA), 2 mM phenyl methyl sulphonyl fluoride (PMSF), 30% glycerol) and the supernatants analysed by SDS-PAGE[31] and fluorography. Lane 1, the initial ^{35}S-methionine-labelled parasite preparation; lane 2, parasite proteins bound to and eluted from the anti-transferrin Ig-beads. The most prominent protein obtained has a molecular weight of 93K. Lane 4, parasite proteins that bind to the transferrin beads. The 93K protein is the most prominent protein binding to transferrin. Though it cannot be seen directly in this autoradiograph, the large amount of transferrin in lane 4 produced a displacement of other radiolabelled proteins in the 70K M_r (open arrow) region. The presence of transferrin was confirmed in Coomassie blue stain of this same gel. Lanes 3, 5 control affinity columns. Although there is some non-specific binding of parasite proteins the 93K protein is not present. Both affinity procedures identify a major protein band of M_r 93K (solid arrow) which binds to transferrin as the parasite transferrin receptor.

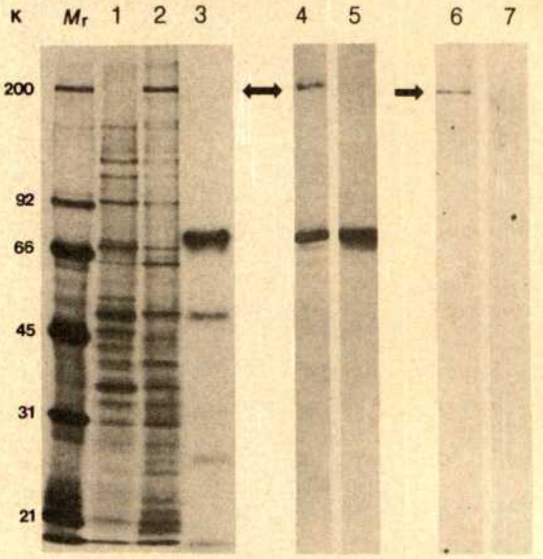


Fig. 4 Incorporation of ^{125}I-Fe-transferrin by *P. falciparum* in culture. Fluororadiograph of reduced SDS-PAGE of *P. falciparum* parasite preparations. *P. falciparum* trophozoite-infected erythrocytes were cultured in complete medium with 7 μCi ml^{-1} of ^{35}S-methionine (Amersham, 332 MBg ml^{-1}) for a total period of 4 h. After 3 h in culture, the preparation was divided into two aliquots, (samples 1, 2). A third was cultured for 3 h without ^{35}S-methionine. ^{125}I-labelled Fe transferrin was added to samples 2 and 3. The three cultures were incubated for a further hour at 37 °C. The parasites were harvested and washed three times with cold RPMI 1640. The parasitized red cells from samples 1 and 2 were lysed with 0.01 M Tris/HCl, *p*H 7.4 and spun at 10,000 r.p.m. for 30 min. This lyses the erythrocyte plasma membrane but leaves the intact parasites surrounded by the parasitophorous vacuole. The parasites were washed with 0.1 M Tris HCl *p*H 7.4 and extracts were prepared by adding solubilizing buffer (10 mM HEPES, 2.0% TRITON-X100, 2 mM PMSF, 5 mM L-1-tosylamide-2-phenylethyl chloromethyl ketone (TPCK), 5 mM *N*-α-p-tosyl-L-lysine chloromethyl ketone (TLCK), 50 μg ml^{-1} leupeptin, 50 μg ml^{-1} chymostatin). The extract was clarified by centrifugation at 100,000 *g* for 1 h. Supernatants containing extracted proteins were added to sample buffer (as described in Fig. 3 above) and run reduced conditions in SDS-PAGE gels. Lane 1, ^{35}S-methionine-labelled trophozoites from sample 1; lane 2, labelled trophozoites cultured in medium with ^{125}I-transferrin, sample 2; lane 3, ^{125}I-labelled Fe-transferrin. (Breakdown products were sometimes present in SDS gels.) ^{125}I-labelled ferrotransferrin alone was also added to sample 3, (young trophozoites in culture) for 1 h. Intact infected erythrocytes from sample 3 were harvested, washed, and dissolved in solubilizing buffer. The extract was clarified by centrifugation at 100,000 *g* for 1 h. The supernatant with the extracted proteins was added to sample buffer and resolved by SDS-PAGE. A 200K protein is present in both lanes 2 and 4), samples 2 and 3 to which ^{125}I-ferrotransferrin M_r 73K was added. The ^{125}I-ferrotransferrin (73K band, lane 4) is distinct from the ^{125}I-label which is present in a 200K protein band. We cannot at present explain how the 73K transferrin is associated with other molecules in SDS to give a 200K band. (Several procedures using different gel conditions are in progress.) Lane 5, the ^{125}I-labelled transferrin added to cells in lane 4. Arrows, ^{125}I-labelled F-transferrin-parasite protein complex (200K). The presence of transferrin in this 200K protein band was confirmed by immunoblotting the isolated parasite preparation from sample 2 (shown in lane 2) with purified rabbit anti-transferrin antibody and alkaline phosphatase-coupled goat anti-rabbit[32]. The result of this blot (lane 6) indicates that there was transferrin in the 200K protein. A control of non-infected mature erythrocytes cultured in medium with transferrin showed no reaction (lane 7).

We next determined the stage in the parasite's life-cycle where maximum uptake of ^{59}Fe bound to ^{125}I-labelled transferrin took place. By monitoring *P. falciparum*-infected cells from ring to schizont stages we found that the young trophozoite stage is most active, internalizing both radioactive labels which then accumulated in or around the parasite within the red cell. Further, after the erythrocyte membrane was removed from the infected cell by immune lysis, we found that both ^{59}Fe and ^{125}I-transferrin were still associated with the parasites isolated from the host cell.

In an effort to monitor changes in the presence of parasite proteins which might bind to transferrin, young trophozoites of *P. falciparum* were divided into two samples and metabolically labelled for four hours with ^{35}S-methionine. ^{125}I-transferrin was added at the end of three hours to sample 2. The parasites in samples 1 and 2 were collected at the end of four hours, then isolated from their red cells; this process was found to retain much of the parasite transferrin receptor. Solubilized parasite proteins from each sample were resolved by SDS–PAGE (Fig. 4). Lanes 1 and 2 show protein extracts from samples 1 and 2, (young trophozoites separated from the erythrocyte membranes); lane 3 shows the ^{125}I-transferrin used in sample 2. It was not possible to determine major differences in levels or the metabolically labelled protein between samples 1 and 2, but we did observe a 200K band in sample 2, suggesting that the ^{125}I-transferrin added to sample 2 was bound to a parasite protein, presumably the parasite transferrin receptor. This 200K transferrin complex is not present in quantities sufficient to be detectable in samples labelled only with ^{35}S-methionine but it must be different from the 195–210K protein seen in late stage *P. falciparum*-infected cells[19] because that protein does not contain transferrin. To confirm the existence of a 200K band in sample 2, we incubated parasites (sample 3) for one hour with only ^{125}I-ferrotransferrin, after which the erythrocytes were collected intact. When the infected cells in sample 3 were solubilized and resolved by SDS–PAGE, we observed a ^{125}I-labelled 200K protein band distinct from the 73K transferrin band (Fig. 4, lane 4). Lane 5 shows the ^{125}I-transferrin which was added to sample 3. The 200K protein band was not seen when ^{125}I-ferrotransferrin was added to control cultures of uninfected mature erythrocytes.

The presence of transferrin in the 200K band was confirmed by immunoblotting the parasite-solubilized protein preparations with purified anti-transferrin immunoglobulin, then alkaline phosphatase conjugated anti-rabbit antibody (Fig. 4, lane 6). No reaction product was seen in the control normal erythrocytes solubilized and treated in the same way as infected cells (Fig. 4, lane 7). The control red cells were identical to those used in the parasite culture. We pretested these control red cells in several ways to ensure that they did not contain reticulocytes, and used the red cells (blood bank blood) in the *in vitro* culture of *P. falciparum* after two weeks storage at 4 °C. We found that these cells did not take up FITC-transferrin or ^{125}I-transferrin, did not bind anti-transferrin antibodies on Western blots or on fixed smears, and did not immunoprecipitate with antitransferrin antibody. Further, these red cells did not take up ^{35}S-methionine above background levels.

Mammalian transferrin receptors are composed of two disulphide-linked 93K proteins that, in turn, bind two molecules of transferrin[20]. We do not yet know the nature of the parasite transferrin receptor/transferrin association. Based on the fact that a parasite derived transferrin receptor must bind mammalian transferrin and, consequently, may resemble the mammalian counterpart, we believe that the 93K protein synthesized by *P. falciparum* is the parasite's own transferrin receptor. It is, however, also possible that the 50K protein seen in the SDS–PAGE gels of affinity-purified parasite proteins is either a fragment of the 93K protein or perhaps functions together with this protein as a part of the receptor. Polyacrylamide gel electrophoresis of purified infected erythrocyte membranes, under nonreducing conditions, suggests such an association of these proteins (data not shown).

In each of the techniques used, we always found the 93K protein associated with or binding to human ferrotransferrin and not apotransferrin, strongly suggesting that this protein serves as a receptor for transferrin. We are now producing antibody to the 93K protein in mice following the procedure of Tung *et al.*[21]. This antibody will be used to determine whether the 93K protein, together with transferrin, is part of the 200K band. Preliminary Northern blot analysis of genomic *P. falciparum* DNA with the complementary (c)DNA clone of the human transferrin receptor indicates that the parasite has genes that code for this protein and that they are different from the human gene[23].

Several different approaches suggest that a *P. falciparum* parasite-derived molecule has transferrin receptor-like behaviour: (1) this molecule binds ferrotransferrin but not apostransferrin on parasitized erythrocytes; (2) FITC-transferrin uptake can be completely inhibited by an excess of unlabelled transferrin; (3) affinity chromatography with antitransferrin antibody or with transferrin reveals the presence of this specific parasite protein, and (4) the molecular weight of the complex of parasite receptor and transferrin from parasitized red cells is different from the transferrin/transferrin receptor complex described for reticulocytes[15,16]. More importantly, our results indicate that malaria parasites transport proteins to the surface of the infected cells (see also refs 23, 24, 25), and use these proteins to obtain nutrients and other elements from the plasma. The parasite within the erythrocyte is in contact with macromolecules in its extracellular environment and is thus exposed to attack by immunological and pharmacological means.

We thank Drs H. Lodish, C. Mulford, D. Titus and J. Glass for their helpful comments and suggestions on this manuscript. This work was funded by the Rockefeller Foundation and by a grant from the Searle Scholars Program to M.J.

Received 18 February; accepted 28 October 1986.

1. Masawe, A. E. J., Muindi, D. M. & Suai, G. B. R. *Lancet* **ii,** 314-317 (1974).
2. Murray, M. S., Murray, A. B., Murray, M. B. & Murray, C. J. *Br. med. J.* **11,** 1113-1115 (1979).
3. Raventos-Suarez, C., Pollack, S. & Nagel, R. L. *Am. J. trop. Med. Hyg.* **31,** 919-922 (1982).
4. Harvey, P. W. J., Bell, R. G. & Nesheim, M. C. *Infect. Immun.* **50,** 932-934 (1985).
5. Fritsch, G., Treumer, J., Spira, D. T. & Jung, A. *Expl Parasitol.* **60,** 171-174 (1985).
6. Pollack, S. & Fleming, J. *Br. J. Haematol.* **58,** 289-293 (1984).
7. Ball, E. G., McKee, R. W., Anfinsen, G. B., Cruz, W. O. & Geiman, Q. M. *J. biol. Chem.* **175,** 547-571 (1948).
8. Moulder, J. W. & Evans, E. A. *J. biol. Chem.* **164,** 145-157 (1940).
9. Levy, M. R., Siddiqui, W. A. & Chou, S. C. *Nature* **242,** 546-549 (1974).
10. Landow, D., Callahan, Jr., E. W. & Schmid, R. *J. clin. Invest.* **49,** 914-925 (1970).
11. Sherman, I. W. & Tanigoshi, L. *Biochemistry of Parasites* (ed. Slutzley, G. M.) 137-149 (Pergamon, Oxford, 1981).
12. Keilin, D. & Hartree, E. F. *Biochem. J.* **41,** 403-406 (1947).
13. Rudzinska, M. A. & Trager, W. *J. Biophys. biochem. Cytol.* **6,** 103-112 (1952).
14. Deegan, T. & Maegraith, B. G. *Ann. trop. Med. Parasit* **50,** 194-211 (1956).
15. Van Bockxmeer, F. M. & Morgan, E. H. *Biochim. biophys. Acta* **584,** 76-83 (1979).
16. Lodish, H. F. & Small, B. *J. Cell Biol.* **65,** 51-64 (1975).
17. Bleil, J. D. & Bretscher, M. S. *EMBO J.* **1,** 351-355 (1982).
18. Van driel, I. R., Stearne, P. A., Greco, B., Simpson, R. J. & Goding, J. M. *J. Immun.* **133,** 3220-3224 (1984).
19. McBride, J. S., Newbold, C. I. & Anand, R. *J. exp. Med.* **161,** 160-180 (1985).
20. Nunez, M. T., Glass, J. & Cole, E. S. *Biochem. biophys. Acta.* **673,** 137-146 (1981).
21. Tung, A. S., Shyr-te, S. & Nisonoff, A. *J. Immun.* **116,** 676-681 (1976).
22. Rogers, J. & Jungery, M. (manuscript in preparation).
23. Leech, J. H., Barnwell, J. W., Miller, L. H. & Howard, R. J. *J. exp. Med.* **159,** 1567-1575 (1984).
24. Schmidt-Ullrich, R., Wallach, D. F. H. & Lightholder, J. *J. exp. Med.* **150,** 86-99 (1979).
25. Wallach, D. F. H. & Conley, M. *J. molec. Med.* **2,** 119-136 (1977).
26. Trager, W. & Jensen, J. B. *Science* **193,** 673-675 (1976).
27. Reese, R. T., Langreth, S. G. & Trager, W. *Bull. WHO* **57** (Suppl.), 53-61 (1979).
28. Bates, G. W. & Schlaback, M. R. *J. biol. Chem.* **250,** 2177-2181 (1975).
29. Markelonis, G. J. *et al. J. Cell Biol.* **100,** 8-17 (1985).
30. Mason, D. Y., Abdulaziz, Z., Falini, B. & Stein, H. *Immunocytochemistry: Practical Applications in Pathology and Biology* (eds Polak, J. M. & Van Noorden, S.) 113-118 (Wright-PSG, Bristol, 1983).
31. Laemmli, U. K. *Nature* **227,** 680 (1970).
32. Burnette, W. N. *Analyt. Biochem.* **112,** 195-203 (1981).

Arc minute gravitational lenses and cosmic strings

PACZYŃSKI has suggested[1] that several large separation gravitational lens candidates could be explained by the existence of cosmic strings. These lens candidates are quasar pairs with separation between 1 and 4 arc min. Paczyński states that arc minute separation quasar pairs, which are lensed by long straight strings, are a prediction of cosmic string theory. Paczyński's work received a great deal of attention when preliminary evidence for a gravitational lens with a 2.6 arc min separation was published by Turner *et al.*[2]. Because a lens with such a large separation is difficult to explain with unseen conventional objects and requires a large mass-to-light ratio for the lens, this lens candidate has been reported[3] as evidence for a cosmic string. Subsequent observations suggest that the quasar pair is not lensed[4]. We point out here that such large separation lenses are not predicted by cosmic string theories.

The typical separation of the images of an object that is lensed by a cosmic string[5-7] is $4\pi G\mu$ where $G\mu$ is the dimensionless mass parameter of the string. A gravitational lens with a separation of 2.6 arc min corresponds to $G_\mu \approx 6\times10^{-5}$. While it is possible for a string configuration to produce gravitational lenses with image separations $>4\pi G\mu$ with favourable geometry, to increase the image separation by more than a factor of two or three above this value seems to require very unlikely conditions. While it was originally suggested[8-10] that a valve of $G\mu \approx 10^{-3}$-10^{-5} might be useful for galaxy formation, further work has shown that these valves are too high, and probably already excluded by observation. Arc minute separation lenses are thus not a prediction of cosmic string theories. If some of these objects were lensed by long straight strings, it would create other cosmological problems, or suggest a basic misunderstanding of how cosmic strings should behave.

We now list the evidence against valves of $G\mu > 10^{-5}$. Models where galaxies or clusters accrete around loops[11,12] have obtained $G\mu \approx 2\times10^{-6}$. Thus the prediction of cosmic string models is that $G\mu \approx 10^{-6}$ not 10^{-5}. More serious problems for large values of $G\mu$ come from nucleosynthesis and anisotropy of the microwave background radiation. The most stringent limit on $G\mu$ comes from the requirement that the strings do produce so much gravitational radiation as to interfere with the successful primordial nucleosynthesis scenario[13-15]. This bound is stringent, $G\mu < 4\times10^{-6}$, but it can also be avoided by some types of cosmic strings that predominantly radiate massless goldstone bosons rather than gravitational waves[16]. However, these strings are not as attractive as seeds for galaxy formation, nor do they appear in any attractive particle physics models. A completely model-independent limit has been obtained by Kaiser and Stebbins[17] who showed that $G\mu < 10^{-5}$ is required by the observed isotropy of the microwave background. Similar limits come from anisotropy due to string loops and the gravitational waves they produce[18,19]. While estimating the errors on these theoretical limits is difficult, we feel that at least the limit of Kaiser and Stebbins is actually conservative. As they stated, the limit given is for a "minimal model" which, for example, does not include the increase in the r.m.s. temperature fluctuation due to the superposition of strings which will occur in any model. Thus, the limit $G\mu < 10^{-5}$ is quite firm, so the predicted separation of images lensed by a cosmic string is definitely <1 arc min.

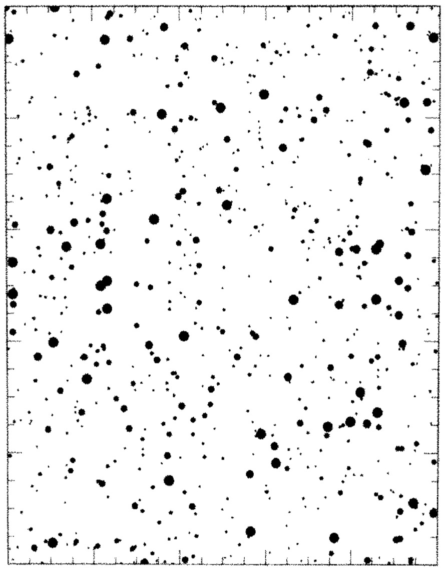


Fig. 1 Computer-generated image of 700 stars in a flat three-dimensional space, with a broad luminosity function.

This work was supported by the Department of Energy, contract DE AC03 76SF00515.

DAVID P. BENNETT
MS 209,
Fermilab,
P.O. Box, 500, Batavia,
Illinois 60510, USA
ALBERT STEBBINS
Physics Department,
University of California, Berkeley,
California 94720, USA

1. Paczyński, P. *Nature* **319**, 567–568 (1986).
2. Turner, E. L. *et al. Nature* **321**, 142–144 (1986).
3. Gott, J. R. *Nature* **321**, 420–421 (1986).
4. Shaver, P. A. & Christiani, S. *Nature* **321**, 585–586 (1986).
5. Vilenkin, A. *Astrophys. J.* **282**, *L*51–53 (1984).
6. Gott, J. R. *Astrophys. J.* **288**, 422–427 (1985).
7. Frieman, J. in *Inner Space Outer Space* (eds Kolb, E. W., Turner, M. S., Lindley, D., Olive, K. & Seckel, D.) 540–542 (University of Chicago Press, 1986).
8. Zel'dovich, Y. B. *Mon. Not. R. Astr. Soc.* **192**, 663–667 (1980).
9. Vilenkin, A. & Shafi, Q. *Phys. Rev. Lett.* **51**, 1716–1719 (1983).
10. Turok, N. *Nucl. Phys.* **B242**, 520–541 (1984).
11. Turok, N. & Brandenberger, R. *Phys. Rev.* **D33**, 2175–2181 (1986).
12. Stebbings, A. *Astrophys. J.* **303**, L21–25 (1986).
13. Davis, R. L. *Phys. Lett.* **161B**, 285–288 (1985).
14. Bennett, D. P. *Phys. Rev.* **D33**, 872–888 (1986) *erratum* (in the press).
15. Bennett, D. P. *Phys. Rev.* (in the press).
16. Davis, R. L. *Phys. Rev.* **D32**, 3172–3177 (1985).
17. Kaiser, N. & Stebbins, A. *Nature* **310**, 391–393 (1984).
18. Brandenberger, R. H. & Turok, N. *Phys. Rev.* **D33**, 2182–2195 (1986).
19. Traschen, J., Turok, N. & Brandenberger, R. *Microwave Anisotropies from Cosmic Strings*, ITP Preprint, NSF-ITP-85-110 (1985).

PACZYŃSKI REPLIES—In my letter[1] I intended to emphasize two issues. Firstly, it is unlikely that cosmic strings could be discovered by finding pairs of images. I pointed out that there is a more striking effect of a string lensing: the images of some galaxies would appear as if they were cut with a sharp edge. Secondly, I pointed out that with the images separated by a few arc minutes the expected difference in the light travel time between the two images may be a few hundred or even a few thousand years. We do not know how stable are the broad line spectra of quasars, and I presented some observational reasons why they may, in fact, be variable. Therefore, I proposed that the forbidden lines in the spectra of pairs of quasars may help to decide between a cluster of quasars and a gravitational lens hypothesis. I quoted some papers (not mine) that suggested that strings may produce pairs of images separated by up to a few arc minutes, but this was just one of the many possibilities I listed. Therefore, it is surprising that Bennett and Stebbins give me so much credit for the cosmic string as the cause of the 2.6 arc min lens candidate.

I would like to take this opportunity to re-emphasize the same two points. Figure 1 shows a computer-generated field with about 700 'stars' randomly placed in a flat three-dimensional space, with a broad luminosity function. Half-way between us and the 'horizon' there is a straight 'cosmic string' that produced 27 pairs of images. Readers are encouraged to find the pairs and the string. This can be done, but the presence of the 'string' is not striking. That was my first point[1].

I do not think that it is possible to prove or disprove the gravitational lens hypothesis for the 2.6 arc min candidate of Turner *et al.* by analysing their spectra, as the forbidden lines are too weak to be seen. Fortunately, the quasars (or the quasar) are relatively close, with redshift, $z = 1.01$. With present technology it is possible to see, although with some difficulty, galaxies at this redshift and beyond. Galaxies have two obvious advantages over quasars for our purpose: their space density is much higher, and they do not vary on a timescale shorter than a few billion years. Therefore, if there is a giant gravitational lens then we should be able to see many multiply-lensed galaxies. If these are discovered we shall know there is a lens, and we shall be able to study the lens structure. If the multiply-lensed galaxies are not found then we shall have to conclude that there is no lens. The task is obviously difficult, and it may take years to complete, but I think that any proof may come only from studies of far away galaxies in the field of the 2.6 arc min lens candidate.

P. Paczyński

Princeton University Observatory,
Peyton Hall, Princeton,
New Jersey 08544, USA

1. Paczyński, P. *Nature* **319**, 567–568 (1986).

'Capture' of stereopsis by illusory contours

Ramachandran and Cavanagh[1] reported an interesting effect in which disparate subjective contours capture regions of zero disparity enclosed within these contours but not regions that lie outside the contours. They suggest that this phenomenon challenges the current computational models of stereo matching. However, the effect is not as general as their article implies: it occurs only in the presence of ambiguity posed by periodic structures. When non-periodic backgrounds are used the stereo capture does not occur (Fig. 1 bottom). This suggests that the phenomenon may have something to do with the treatment of periodic structures by the visual system[2].

Most current computational models of stereo matching[3–5] are based on point-by-point matching of features and use a set of assumptions or constraints (based on some very general properties of the environment) to ease the computational burden associated with solving the often massive local ambiguity problem. Surfaces containing periodic texture cannot be successfully matched locally. For example, a periodic texture on a frontoparallel plane will produce a set of peaks in the cross-correlation surface, each corresponding to a possible match. Locally, there is not enough information to decide which of these possibilities is the correct one. True correspondences can be obtained only by using more global information or by incorporating information originating outside the stereomatching system. This is true for any matching mechanism relying on local computations.

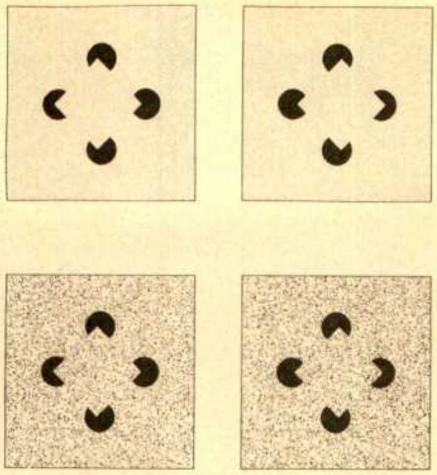


Fig. 1 Depth carried by disparate subjective contours is attributed to periodic (top) but not to non-periodic (bottom) texture elements enclosed by these contours. The stereograms were constructed for viewing with free fusion by crossing the eyes in front of the displays.

The problem of multiple 'local maxima' has been often attacked by invoking multiresolution strategies in which low spatial frequency matches are used to disambiguate the high frequency ones[4]. However, such strategy is not of much use in the case of extended periodic structures. Another, more general, way to solve this problem is to use other cues, for example, accommodation[6], to bias disparity. The idea of a set of depth maps cooperating to converge on a state in which all maps agree and present sharp and accurate three-dimensional localization of features is a powerful notion, the potential of which has yet to be fully explored in the machine vision research.

Current computational theories of stereopsis do not address the qualitative or quantitative aspects of binocular depth perception. Rather, they attempt to understand only its first step: the detection of binocular disparities. Not all of them assume that binocular disparities are directly linked to our perception. That is to say, not all of them assume that our perception of depth from stereograms are isomorphic to the state of the neural disparity network. It is possible, to use an analogy, that binocular disparities are to the stereoscopic depth perception what the output of retinal receptors is to the perception of brightness.

Rather than challenging current computational models of stereomatching, stereoscopic capture illuminates, I believe, mechanisms underlying more cognitive or configurational aspects of stereoscopic processing.

1. Ramachandran, V. S. & Cavanagh, P. *Nature* **317**, 527–530 (1985).
2. Mitchison, G. J. & McKee, S. P. *Nature* **315**, 402–404 (1985).
3. Mayhew, J. E. W. & Frisby, J. P. *Perception* **9**, 69–87 (1980).
4. Marr, D. & Poggio, T. *Proc. R. Soc.* **B204**, 301–328 (1979).
5. Prazdny, K. *Biol. Cyb.* **52**, 93–99 (1985).
6. House, D. in *Proc. Visuomotor Coordination in Frog and Toad Workshop* (ed. Arbib, M.) Tech. Rep. 82-16 (University of Massachusetts, Amherst, 1982).

K. Prazdny

Machine Perception Group,
Artificial Intelligence Laboratory,
FMC Corporation,
1185 Coleman Avenue,
Box 580, Santa Clara,
California 95052, USA

Ramachandran replies—in our article on stereoscopic capture[1] we argued that existing stereo-algorithms could not account for our findings although it is possible that some of them could be modified to do so. This is quite consistent with Prazdny's interesting suggestion. Since disparity of illusory contours can influence the matching of finer elements in a stereogram (Fig. 1, top), we concluded that monocular contours are involved in stereopsis[1,2]. Contrary to Prazdny's suggestion, however, we believe this to be a very general effect and 'capture' happens to be just one way of demonstrating their role. Prazdny also suggests that the effect is specific to 'periodic' patterns, a possibility that we ourselves consider (page 528) when discussing the phase sensitivity of stereo-capture. However, it is almost certainly the lack of ambiguity (not lack of periodicity) that prevents capture in Prazdny's Fig. 1 (lower).

Furthermore, there are also three other aspects of our effect that are not easily handled by existing computational models. First, stereo capture can be seen even in rivalrous displays[1,2] (especially when flashed briefly) and not just in peri-

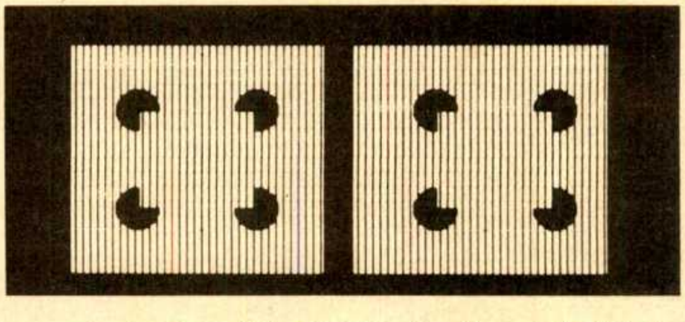


Fig. 2 Stereo capture is not seen if disparity is introduced between the disks themselves.

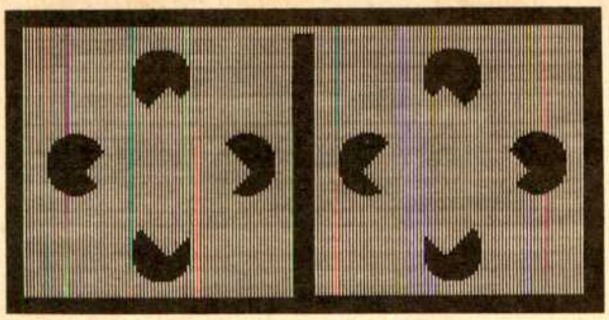


Fig. 3 Occlusion must be an inherent part of the stereoscopic matching process. The four black disks now become 'holes' through which one can see the four deep corners of a partially hidden diamond and the corners pull the corresponding lines of wallpaper with them.

odic displays. Second, Prazdny suggests that disparity signals from the 'unambiguously matched borders' help obtain true correspondences. This conjecture cannot be valid, at least in its simple form, as Fig. 2 does not yield stereo capture. (Here the disparity is between the disks themselves and not just between the cut sectors.) Hence construction of a three-dimensional surface and the presence of a steep disparity gradient are prerequisites; the mere propagation of depth signals will not suffice. Third, we find that reversing the sign of disparity does not simply reverse perceived depth. When the pictures of the two eyes are interchanged to convey uncrossed disparities (Fig. 3), a qualitatively different percept emerges. The black disks now look like holes through which one can see the four corners of a partially occluded square (diamond) and these deeper corners pull the corresponding lines with them. The illusory contours are now seen to complete the holes rather than the square. Here, the illusory contours are created after stereo matching and must, in their turn, influence the subsequent stereo matching of finer elements.

Effects similar to stereo capture can also be observed in apparent motion (motion capture)[3] but the effect is not phase sensitive and has a much higher tolerance for uncorrelated textures.

Any visual process can in principle be simulated in a computer but our results suggest that the brain solves the correspondence problem (in motion and stereopsis) by using a set of short cuts[4] rather than elaborate computation of the kind implied by artificial intelligence researchers. During phylogeny the visual system may have acquired many such tricks which were adopted not for their aesthetic appeal, versatility or mathematical elegance but simply because they worked—an idea that we call the utilitarian theory of perception[5]. Artificial intelligence can be useful in rigorously reformulating[6,7] perceptual problems, but the mechanisms used by the brain to solve these problems are best revealed by old-fashioned psychophysics and neuroanatomy[5,8]. Capture probably results from the fact that the motion (or depth) of certain salient image features is first extracted by the magnocellular pathways and then attributed to finer features that excite the parvocellular pathways. The nature of this interaction merits further attention.

V. S. RAMACHANDRAN

Department of Psychology C-009,
University of California at San Diego,
La Jolla, California 92093, USA.

1. Ramachandran, V. S. & Cavanagh, P. *Nature* **317**, 527–530 (1985).
2. Ramachandran, V. S., Rao, V. M. & Vidyasgar, T. R. *Nature* **242**, 412–414 (1973).
3. Ramachandran, V. S. & Cavanagh, P. *Vis. Res.* (in the press).
4. Ramachandran, V. S. & Anstis, S. M. *Scient. Am.* **254**, 102–110 (1986).
5. Ramachandran, V. S. *Perception* **14**, 97–103 (1985).
6. Marr, D. *Vision* (Freeman, San Francisco, 1982).
7. Ullman, S. *The Interpretation of Visual Motion* (MIT, Cambridge, Massachusetts, 1979).
8. Livingstone, M. S. & Hubel, D. H. *J. Neurosci.* **4**, 309–356 (1984).

Is the Gulf of Tadjoura a platelet boundary or complex rift zone?

CHOUKROUNE *et al.*[1] have provided new data for the structure of the Gulf of Tadjoura (central Afar), which supposedly separates the Danakil and Somali platelets. Their observations (orientation and activity of a NW–SE fault set) made during dives with the *Cyana* submersible, are at odds with the model of orthodox plate-tectonics boundary, which had been developed in the central Afar and in the Gulf, and was previously widely accepted. Choukroune *et al.* question the westwards opening of the Gulf by rifting: their new interpretation of the Arabia and Africa plate-tectonics relationship does not imply that the Afar triangle is a multi-platelet building, but that it could represent a diffuse relay zone between the Aden Rift and the Red Sea Rift.

The new data support observations (unpublished) made during earlier cruises by Boucarut *et al.* Their complete tectono-bathymetric chart of the Gulf of Tadjoura shows that the Maskali–Obock trough and the Tadjoura trough are located on either side of a subcontinuous tectonic lineament (major fault zone), running ENE–WSW (N 60°) from 43°20 to 43°. Both troughs are delineated either by ENE–WSW, NW–SE, or EW fault zones, which form some important (Recent) submarine cliffs. Moreover, the presence of active (NW trending) extensional fault zones, some of them crossing the Gulf and cropping out on both sides[2], and of faulted Middle Pleistocene sediments in the Obock trough and on the northern side of the Gulf[3] had been mentioned.

Continuation of the major structures on land into the Gulf suggests that the Gulf floor is composed mainly of broken continental crust, separated by relatively narrow and irregular shaped rift zones filled with young volcanics[2].

The various orientations of the fault sets in the Gulf should be compared with those of the nearby land area, whose successive motions and combinations with volcanic occurrences during Plio-Pleistocene times are well known[4]. On land, the structural dominance of the 'Aden type' fractures (N 80°–100°) diminishes westwards: likewise, in the Gulf, it is taken over by that of the NW–SE and ENE–WSW faults. This explains the zigzag shape for the (supposed) boundary.

So, even if the platelet sketch for the central Afar has to be reviewed, various data indicate that the westwards tearing interpretation of the Gulf of Tadjoura[5–7] should not be eliminated.

M. BOUCARUT
M. CLIN

University of Bordeaux 3,
33405 Talence Cedex, France

1. Choukroune, P. *et al. Nature* **319**, 396–399 (1986).
2. Boucarut, M., Clin, M. & Dars, R. *C.r. hebd. Séanc. Acad. Sci. Paris* **285D**, 491–494 (1977).
3. Clin, M. & Pelissier-Hermitte, G. *C.r. Campagne Orgon IV, CEPM-CNEXO*, 183–188 (CNRS, Paris 1979).
4. Boucarut, M., Clin, M., Pouchan, P. & Thibault, C. *Geol. Rdsch.* **74**, 123–137 (1985).
5. Boucarut, M. & Clin, M. *C.r. hebd. Séanc. Acad. Sci. Paris* **282D**, 1145–1148 (1976).
6. Courtillot, V. *Phys. Earth planet. Inter.* **21**, 343–350 (1980).
7. Boucarut, M. & Clin, M. *Atti Convegni Lincei, Roma* **47**, 541–554 (1980).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

nature

Europe runs before it walks

The European Commission threatens dire consequences if it is not given a bigger budget for research. But its case is fragile, at best

THE European Commission seems to be heading for a fruitless confrontation with its sponsors, the twelve governments now members of the European Communities. Although this will not be the first time that the Commission has been at loggerheads with the governments of which it is the titular servant, and although there will be general sympathy with the Commission's frustration at the slow pace of change in Europe, it is a pity that the Commission has picked on the research budget on which to do battle. The difficulty is that the Commission is talking as if it might tell member governments that if it cannot have $8,000 million (over five years) to spend on its "framework programme" of largely applied research, it will let the European research enterprise go hang (see *Nature* **324**, 291; 1986). European researchers will be the chief losers, but the Commission's own European ideals will also take a bashing. The best hope now is that the Commission is merely bluffing. Otherwise, it should think again.

The nature of the row is best understood and is only understandable in terms of the Commission's own constitution. Those who wrote the Treaty of Rome which legitimized the European Economic Community and Euratom (there was previously the Iron, Steel and Coal Community) believed that their ultimate goal of a united Europe would come about only as member governments were given proof that Europe would work in small ways. Ingeniously, the founding fathers legislated for a kind of government-in-waiting (the Commission) whose immediate function would be to formulate policies for approval by member governments. In the short run (which is now), the Commission would propose and member governments would dispose, by the agency of what is called the European Council. But in the long run, the calculation ran, member governments would be so content with the developing idea of Europe that they would cede more and more executive power to the Commission and would use the European Parliament as a common legislative body that would also breathe down the Commission's neck. This dream is still a very long way from fruition. The Commission, understandably, is frustrated at being kept a dogsbody for quarrelsome governments, and this frustration has now spilled over onto its policy for research.

Discontents

If discontent with the snail's pace of change were the only touchstone, the Commission would be surrounded by friends. One prominent companion in discontent, but not necessarily an ally, is President François Mitterrand of France, who has been pressing for the past five years for more technical collaboration within Europe and, for the past eighteen months, for a substantial move towards European integration on a wider front. The Commission now seems to regard President Mitterrand's more recent initiative as a licence for insisting that member governments must at last be forced to collaborate on something, and has picked out research. For the time being, it seems prepared to overlook the more glaring of the defects of its enabling treaty — the assumption that little would happen in foreign affairs where, in the event, members' interests most often coincide and the legal provisions that compel collaboration in agriculture to the great cost of European consumers and cheaper food producers elsewhere in the world. Two-thirds of Europe's energy in the past 30 years, and a great deal of mutual recrimination among the Communities' members, have been occasioned by the unkeepable promise to 1950s' farmers that they would be protected from the discomforts of clearly impending technical change. The Commission knows (it has tried) that reforming the Common Agricultural Policy would do more to liberate European energies than a coordinated policy on applied research, however well conceived the programme. Yet the Commission is so exercised about the budget for its framework programme that M. Jacques Delors has thrown his reputation behind the campaign for $8,000 million.

Misunderstanding

The Commission's pitch for a larger research programme would be more telling if its record on the conduct of scientific affairs were more convincing. Unfortunately, its record is poor, to say the best of it. Throughout the long argument on regulations on environmental pollution in the 1970s, the Commission too obdurately clung to the notion that pollution must be controlled by emission standards, so as to be fair to manufacturers of the same kind wherever they might be located, and obscurantistly resistant to the notion that what matters is the degree to which the environment is polluted. On many occasions, it seemed that the difficulty was the Commission's failure to understand the difference. More recently, the Commission has been responsible for two quite shocking assaults on good sense in its handling of essentially technical issues. Thus, in the days after the nuclear accident at Chernobyl in April, the Commission hurriedly organized a meeting to agree on common standards of when it would be prudent to ban radioactively contaminated food and concluded that food should not be sold with more than 1,000 Bq per kg—a standard roughly an order of magnitude more stringent than necessary. Worse still, the Commission has decided to use its powers for the regulation of agricultural produce to ban the use by European farmers of natural growth hormones (bovine growth hormone, perhaps manufactured synthetically, to make more beef, for example). It is generally agreed that such a practice would damage consumers, the people who might eat meat grown in such a way. The Commission's motives have simply been to restrict the embarrassing productivity of European farmers. When the ban was declared (it comes into effect this month) the commissioner responsible for agriculture said that there were occasions when "politics must take precedence over science". Is that the kind of organization to which people will willingly entrust $8,000 million for research?

The Commission will no doubt complain that these administrative matters have nothing to do with its reputation for the conduct of research, but even that is only half true. In its administration of the Communities' four research centres, the Commission has been over-sensitive to political (especially Italian) objections to the suggestion that some of them might be closed. The Commission's sponsorship of fusion research, on the other hand, is admirable. There is no suggestion that the major programmes in applied science have been tainted by political considerations, but the most interesting of them, that called ESPRIT (where the last two characters stand for information technology) is too young for its benefits to be properly evaluated and costed. Most probably, the Commission is right to think that there are good works it might perform in fields like this, but little would be lost and much might be gained if it set out to foster collaboration in research with much more deliberation.

This is not to say that the Commission must forever drag its feet on research. Curiously, there are many causes that the Commission has itself in the past espoused that could benefit from encouragement — and a little money. Only three weeks ago, the annual general assembly of the European Science Foundation was wringing its hands over the poverty of its plan to build 'networks' of European scientists, ironically stimulated by a study sponsored by the Commission a few years ago. The Commission could also usefully (as it may yet) listen to appeals such as that from the group of European radioastronomers who seek a modest 16 million European Currency Units (roughly £11.4 million) over five years to establish a centre for the analysis of data gathered during very-long baseline measurements. Such projects are not, of course, alternatives to a programme of applied research, but the fact that they go begging is a sign that the Commission has not yet begun to work effectively on the common ground where success would come most easily and bring the greatest benefits — the effective integration of the technical community without which a common research enterprise will have little purpose. □

Licensing the spooks

The British and US governments are both embarrassed by the doings of their undercover people.

WHO would have known, until the British government's attempt to prevent in the Australian courts the publication of a book on its security service that Lord Rothschild, the embryologist and merchant banker, was once a member of that same service, called MI5? This is nevertheless what has emerged from the embarrassing (for the British government) disclosure that Rothschild acted as the intermediary in bringing together Mr Peter Wright, once also a member of MI5 and now living in Australia, and Mr Chapman Pincher, who published a racy account of Wright's tale three years ago. Few will think the worse of Rothschild now; some will now put him in the same daredevil category as Colonel Oliver North, the staff member of the US National Security Agency dismissed last week for alleged gun-running (to Iran) and now described by President Reagan as a national hero.

Popular heroics apart, there is nevertheless a case for believing that spying and spy-catching have got out of hand. Every few months, there seems to be either a serious international incident or a serious domestic crisis in some major capital because of spooky doings (Auckland Harbour, for example). Then there are retaliations (the name of Daniloff refers). The obvious difficulty is that espionage and counter-espionage are unregulated. The obvious solution is that governments employing secret agents should be required to register all salaried agents' identities with an international authority, which would keep the information secret unless asked by the registrant to disclose it (perhaps secretly) to a third party. That way, it would be possible for governments more easily to arrange swaps for agents or even to find out whether people arrested in foreign capitals were in their employ or were merely innocent visitors. By becoming licensed in this way, secret agents would be able to trade self-esteem for the dare-devilry that now sustains them, and would be less inclined to write their memoirs. So long as the international authority were not itself penetrated by agents, the whole business would be made less serious a threat to international good relations; international spy crises might, for example, be defused by mutual consent. Would not such a scheme be worth a trial? □

Action against AIDS

Governments are moving quickly on publicity about AIDS; should they be doing more?

BOTH the British and the French governments have responded well to the latest alarm about the spread of AIDS (acquired immune deficiency syndrome). In both countries, the governments have decided that public education must take precedence in any serious attempt to break the chain of infection by the virus. Plainly that makes sense. Unless people know which features of their personal behaviour expose them to the risk of infection, their self-interest will not coincide with the needs of programmes of public health aimed at averting what could be a cataclysmic infection. One of the several haunting aspects of the problem is that there remain so many features of the natural history of AIDS that are, as yet, unknown. It is not yet known when and for how long, in the course of an infection, a person is capable of infecting others. Whether all those infected, or only a proportion of them, eventually develop AIDS, is another imponderable for the time being. Little is known about the relative risks of transmission in different sexual acts. The case for acting now is that, if all the answers to such questions were gloomy, large proportions of us or our successors would be killed.

It is a pity that young people, at the beginning of their sexual lives, must be the chief targets of public education, but that, too, makes sense; it is not, as usually the case, that young people are especially deserving of consideration (true though that remains) but that they could be especially effective agents for the transmission of a socially damaging and usually venereal disease. To say the least, a young person, during the remainder of his or her lifetime, will on the average have more sexual partners than an older person and will therefore, on the average, be a potential source of infection for a greater number of people. It is also possible that some young people may be unusually promiscuous. To their credit, both governments appear to have faced this part of the problem that confronts them squarely. The British government, with its reputation for being careful about money, has surprised its taxpayers by deciding to spend twice the sum it first thought of (£20 rather than £10 million) on immediate publicity. The French government is wisely emphasizing the need that people should have a ready opportunity of a test for antibodies.

None of this, unfortunately, will throw light on the present incidence of the infection (as distinct from the disease it may ultimately cause). Those who voluntarily seek an antibody test will often be people who suspect they may have run a particular risk. The routine testing of blood offered at transfusion clinics may similarly, as in the United States, be biased by people's recognition that a blood transfusion is the most convenient way of being tested. That is why some means of random testing would, in present circumstances, be beneficial; it would then be possible to tell objectively how great is the present pool of infection, a necessary starting point in estimating more accurately the rate at which infected people progress to overt disease. Schemes for the anonymous testing of pregnant women at antenatal clinics have similarly been advocated as ways of monitoring the spread of AIDS through heterosexual intercourse. The snag is that, in these enlightened days, it is unethical for physicians to test a person's blood without telling him or her what is afoot, whereupon it would be unethical not to disclose the results. The result is that information that would be invaluable for an assessment of the present problem, and which could be collected easily, is not available. The reasons for this self-denial are admirable, but the outcome is ignorance on an important point. If the AIDS problem becomes more serious, societies will find themselves considering a variety of illiberal measures to protect themselves against the pool of those infected; on the face of things, anonymous testing whose results are not disclosed might be a lesser evil that should be more quickly allowed. □

French universities

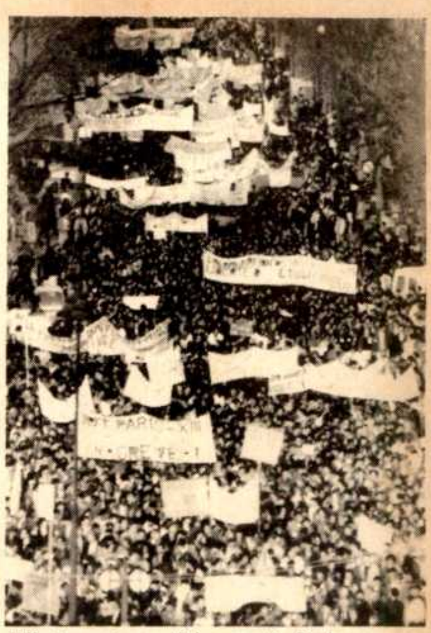
Student protests succeed in postponing new law

For 1986 read 1968 — the year of revolution in French universities which led finally to the resignation of General de Gaulle as President of France. Last week, more than half a million French students and schoolchildren were on the streets to protest against a new law for the universities which, students claim, will tighten up entry requirements, reverse the right of free entry to universities that was won so bitterly by, in many cases, the demonstrators' parents nearly 20 years earlier. Right-wing Prime Minister Jacques Chirac, who dealt with the events in 1968 as a protégé of de Gaulle's prime minister and later president, Georges Pompidou, this time had no stomach for a fight. He told his minister for the universities and research, Alain Devaquet to withdraw the bill for revision.

Resignations do not seem to be in order this time, however. Chirac is not going to see the fall of himself or his government over the issue when the national budget and the economy is his main theme; and the socialist President, François Mitterrand, who might be contending with Chirac in the presidential elections of 1988 and had no involvement in the preparation of the bill, must be laughing up his sleeve. Furthermore, Devaquet, whose law was a compromise between various pressure groups, is said to have many sympathies with the students, and may see the demonstrations as giving weight to his own position. The losers so far in the battle would appear to be *'les emigrés'*, that group of university staff which was silent, or powerless, during the previous socialist administration but which now that a right-wing government is in power has been calling for much greater power for the university professors in France, both in the selection of students and over the control of research funds. In the latter case this would have entailed the dismemberment of the French national research council, CNRS, a battle the *emigrés* recently lost; now, it seems, their battle for more control over their students, 40 per cent of whom leave after one year in university, has also gone against them.

There is thus a good deal of despondency among the rightist camps in French universities this week — if this government cannot help them none can. On the other hand, French natural scientists, who number few of the *emigrés* among them, are more cheerful. Many deplored yet another attempt to reform the universities, and most thought the new bill's efforts to divide departments into single-subject faculties retrogressive, likely to reduce interdisciplinary research and to recreate the pre-1968 professional fiefdoms.

Another common complaint of researchers about the new bill was that it threw away the fledgling French PhD, just two years old, to put back the old system of multiple theses and doctorates of various levels which was incompatible with research qualifications outside France and reduced international mobility. The bill before parliament last week, which resulted in the student demonstrations, had been improved in this respect by amendments in the French Senate, but many researchers are still unhappy.



Associated Press.

Effective yet peaceful protest in Paris.

ADEMAST, the Association Nationale Pour le Développment et la Maîtrise des Sciences et des Techniques, an association of scientists that grew up after the epoch-making 1982 national colloquium on science and technology in Paris has also strongly criticized the bill, now withdrawn after the protest. In a report published just before last week's demonstrations, ADEMAST admitted deficiencies in the universities but rejected the constant politicization of university reform.

Governments "must work with what exists" rather than attempt in each parliament to recreate the universities anew, said the scientists' associates. And it gave a strong endorsement to the system of multiannual contracts, by which the ministry of education in the previous administration was guaranteeing extra funds to universities.

These funds were available only to establishments which put in place a serious research and training policy to meet local and national needs. These multiannual contracts changed the distribution of university funds from what had been uniform "sprinkling" to real selectivity, the association claimed. But the new bill appeared to be weakening this kind of selection rather than strengthening it.

The scientists' association recommends the creation of an agency to distribute university funds selectively. The agency would be indepedent of the ministry, and similar to the British University Grants Committee; but again the new law moved in the opposite direction in the name of autonomy. According to ADEMAST members, French universities are still immature after the revolution of 1968; and the new law would have led to a mere counter-revolution rather than an improvement. Whether any changes at all will now occur is yet to be seen.

Robert Walgate

Hermes to win approval soon?

Britain's space industry was earlier this weeks still awaiting government approval of a proposed multimillion dollar contribution to the European space shuttle project, Hermes. The approval, due last Monday (1 December), is for about 10 per cent, or $4 million, for the first phase of the space project. The preparatory phase, which began in the autumn and will take a year to complete, will cost $40 million. The complete programme, geared to a launch in 1995–96, will cost $1,200 million, of which France and West Germany will provide 39 and 30 per cent, respectively.

But a great deal of research is needed in the preparatory phase to tell how companies with expertise in space technology will benefit. Nearly 30 submissions on how the technology can be advanced and exploited are with the European Space Agency, which is coordinating the project.

The British National Space Centre, formed about a year ago to coordinate the British space effort, awaits a response from the government to its 15-year plan, which includes the Hermes project. The centre wants an early decision so that preparations can be made for a ministerial meeting of agency members next June.

The Hermes is a space aeroplane to be launched aboard the new European rocket Ariane V, which will be ready for its first flight in about 1994–95. The research on Hermes is expected to inspire applications in other military and civil aviation fields. New types of wings and flight systems would be two possible areas of research.

Bill Johnstone

Chemical weapons

Study reflects Atlantic divisions

Washington

EUROPEAN and US perceptions of the value of North Atlantic Treaty Organization (NATO) chemical weapons are still very much at odds, to judge from a new joint study* by independent strategists. The Aspen Strategy Group in the United States and the European Strategy Group failed to find common ground on the key question of whether NATO's policy of modernizing chemical weapons while removing them from Europe is justifiable. And both are pessimistic about a total ban.

Chemical weapons remain a source of deep divisions within the NATO alliance; although European countries want a treaty banning them, the United States is keen to preserve the chemical option as a means of countering Soviet first use and to decrease reliance on nuclear weapons.

Under NATO rules of engagement, chemical weapons would be used only if Warsaw Pact countries used them first. The Soviet Union has more advanced chemical capabilities than NATO, and opinions differ over their importance to Soviet war plans.

The Aspen/European study believes that chemical weapons are probably not a Soviet "weapon of choice". But it could not agree on the wisdom of the planned modernization of the NATO chemical weapons stockpile.

Many, not least the US government's General Accounting Office, have serious reservations about the likely usefulness of the 'Bigeye' binary chemical bomb now being developed (see *Nature* **321**, 717; 1986). Binary weapons are in principle safer than the single-agent weapons now in storage, because two relatively non-toxic agents do not become lethal until mixed during flight. But the Bigeye has had severe technical problems.

Unable to reach joint agreement on NATO strategy, the Aspen/European study settled instead on the face-saving formula that chemical weapons should not be a cause of alliance disunity in deterring conventional or nuclear war in Europe. The group almost tries to have it both ways by saying there is no justification for NATO acquiring a large-scale chemical war-fighting capability, but also that few believe NATO policy should allow nuclear deterrence to be the only means of deterring possible Soviet use.

The compromise seems to be that a small offensive capability would be prudent, which is what NATO has. Given that, the group could find no rational reason for removing them from Europe to the United States. The NATO plan to airlift chemical weapons across the Atlantic in a time of crisis is seen by the group as substantially limiting their deterrent value; the logistical burden in a military crisis would be very substantial.

Negotiations for a comprehensive chemical weapons ban have been in progress for nine years, with little sign yet of agreement. Verification is, of course, the stumbling block, with the United States and the Soviet Union deadlocked mainly over the question of on-site inspections.

The Soviet Union has accepted the principle that destruction of chemical weapons stocks should be monitored continually by on-site inspectors, and appears to be moving towards accepting the need for a quota of on-site inspections to verify compliance. But the United States recently hardened its position, arguing the need for a short notice (48-hour) mandatory challenge procedure.

The Soviets so far reject anything but voluntary inspections, although there is a British proposal that might break the deadlock. Under the proposal, an on-site inspection could be legally refused if the accused party provided evidence that the inspection was unjustified. The proposal fails, however, to specify the type and standard of "evidence of innocence" that would suffice. **Tim Beardsley**

---

* *Chemical Weapons and Western Security Policy*, Aspen Strategy Group and European Strategy Group, 1986.

National Science Foundation

New centres for biotechnology

Washington

THE US National Science Foundation (NSF) plans to spend more money on shared instrumentation and multidisciplinary basic research related to biotechnology. The new programme has two strands — facilities centres to be set up in 1987 and research centres to follow a year later.

Eight million dollars have been appropriated for the development of multi-user instrumentation facilities, or 'mini-centres', in 1987. The foundation expects to make 10–15 two-year awards averaging $500,000 each for these centres, which are being developed to meet the demands of the "consistent lobby of people [doing basic biological research] who need access to sophisticated instruments for less than full-time use", according to Thomas Quarles at NSF.

The instruments NSF has in mind include costly items such as DNA or protein synthesizers, peptide sequencers and the new nucleotide sequencers developed earlier this year. Prices vary, with the DNA machines costing between $30,000 and $90,000, and the peptide machines ranging between $85,000 and $160,000. These prices are prohibitive for many research laboratories, especially if they need the machines only occasionally.

The instrumentation sharing programme has two objectives. The first is to maximize the return on investment by attempting to ensure nearly full-time use of expensive equipment by combined groups of researchers. The second is to stimulate multidisciplinary exchanges through the common use of equipment.

The centres are not, however, designed to be regional facilities located equidistant from the researchers who will use the instruments. David Kingsbury, the assistant director for the directorate administering the programme at NSF, says he sees the instrumentation awards in 1987 going to single departments or groups of investigators in one area.

According to Kingsbury, the success of previous regional shared instrumentation centres has been mixed, depending on the users' location. Often, the institution physically housing the equipment has received the greatest benefit, with little regular sharing with researchers elsewhere. To qualify for support under the research centre scheme in 1988, laboratories will need to demonstrate the quality of their research programme as well as their need for equipment. The aim is to encourage people from various disciplines to work together on important problems in biotechnology. The money is expected to establish, modify or support existing multidisciplinary biological research centres, and the grants will be structured to fill gaps not financed by industry, universities or other government entities.

Kingsbury expects "everyone under the Sun" to apply for the research centre grants, and foresees joint proposals from, for example, ecologists and microbiologists and from chemists and biologists. These awards will depend on the treatment by Congress of the NSF budget for 1988, not yet published, but may be as large as $2–4 million per centre.

NSF says that it has no preconceptions of the topics these centres should deal with except that they should be within the field of biotechnology and should not be related to diseases or the development of drugs, which are the responsibility of other agencies.

Proposals for the 1987 facilities centres are due by 1 April 1987, and those for the 1988 research centres by 1 August 1987. Awards will be made after the deciding panel meets in May or June. James Brown, the division director at NSF, says it will be "nip and tuck" to make the facilities awards within the 1987 fiscal year, which ends next September. **Carol Ezzell**

Telecommunications

Fight for Japanese network

Tokyo

As Tokyo is one of the three principal financial centres of the world and growing the world's telecommunications giants are fighting to plug their cables into Japan. But an attempt by Britain's Cable and Wirelss to tap Japan's lucrative international telecommunications market has met with a solid wall of resistance.

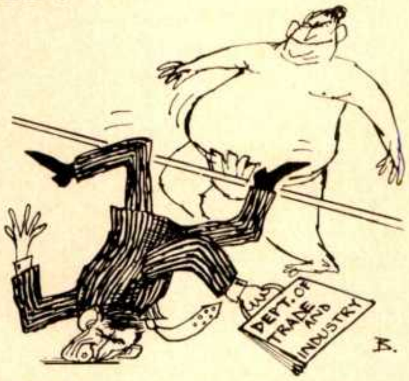
Following last year's liberalization of Japan's telecommunications law, a consortium headed by Cable and Wireless and C. Itoh, a major Japanese trading house, declared its intention to set up an international telecommunications company that would break the monopoly held by Kokusai Denshin Denva (KDD).

But Mitsubishi Corporation, Mitsui and Co. and Sumitomo Corporation beat the C. Itoh/Cable and Wireless consortium to establish a feasibility study company, International Telecom Japan (ITJ), in July (see *Nature* **321**, 803; 1986).

Whereas ITJ plans to follow established telecommunications routes — Intelstat satellites and a submarine trans-Pacific cable (TPC-3) due to be laid by KDD and American Telephone and Telegraph in 1988 — the C. Itoh/Cable and Wireless group was expected to link up with Cable and Wireless' global digital highway, a network of submarine and transcontinental optical fibre cables connecting Western Europe, the United States and the Far East. But there has been strong opposition to the C. Itoh/Cable and Wireless plan.

Both the Ministry of Post and Telecommunications and KDD maintain that there is not enough room for a third company, and KDD says there is no need for an extra trans-Pacific submarine cable. But the most conservative estimates of Japan's international telecommunications market forecast a tripling in size to Y600,000 million (£2,600 million) by the mid-1990s.

Taking heed of the ministry's pronouncements, the C. Itoh/Cable and Wireless group approached the major shareholders of ITJ in October and suggested a merger but their offer was turned down. And in mid-November the group formed International Digital Communications Planning Co. (IDC) with Cable and Wireless and C. Itoh the major shareholders. No sooner had staff set up its new office than Japan's minister of Post and Telecommunications, Shunjiro Karasawa, let drop a bombshell. In discussions with British Trade Secretary Paul Channon last week, Karasawa announced that although Cable and Wireless could keep its 20 per cent share in IDC (under Japanese law a foreign company can hold up to 33 per cent of a telecommunications company), the British company would be barred from management decisions "because international communications are relevant to national security in emergency situations" and thus a foreign company could not have an influential voice or power in a "second KDD".


DEPT. OF TRADE AND INDUSTRY
B.

With ITJ expected to apply for a licence from the Ministry of Post and Telecommunications by the end of this month, Karasawa's statement could not have come at a worse time for IDC. And although one of IDC's managing directors insists that the matter is by no means closed, he also emphasized that Cable and Wireless's trans-Pacific cable is only one of the options currently under consideration by his company. **David Swinbanks**

India creates biosphere reserve

Bangalore

INDIA's first biosphere reserve, patterned after the guidelines in UNESCO's Man and Biosphere Programme, has been inaugurated at Nilgiris in South India, one of twelve biosphere reserves suggested by the Department of Environment (DOE).

It covers an area of 5,520 km^2 comprising the famous Silent Valley, a region with 1,500 plant species and virgin forests of western ghats spanning the states of Kerala, Tamil Nadu and Karnataka. The Nilgiris biosphere reserve, classified as a Malabar rain forest, includes 2,020 km^2 of minimally disturbed forests in which about 20 tribal groups have been identified.

The reserve is cooperatively managed by the three states under existing state and central legislation on conservation and preservation of forests, wildlife and environment. India has 53 national parks and 247 wildlife sanctuaries covering 100,000 km^2.

According to Mr Z.A. Ansari, Union Minister of State for Environment, a comprehensive national conservation strategy is being formulated and steps are being taken to improve the living conditions of tribals in the biosphere reserves.

Radhakrishna Rao

Hepatitis B

Cheap vaccine for wide use?

Washington

AN international "task force on Hepatitis B immunization" has announced plans to use a new vaccine costing as little as $1 per dose in mass immunization demonstration projects in four developing countries. The nine-member task force, which includes several recognized authorities on the disease, aims to convince governments of countries where the disease is endemic that mass immunization, previously considered too expensive, is economically feasible.

Hepatitis B is most common in Asian countries. It causes 80 per cent of the world's liver cancer, and may be second only to tobacco as a cause of world cancer mortality, acccording to one task force member, Dr Richard Mahoney. As little as a year ago, a dose of vaccine cost $30, with three doses needed for a protective course, but Cheil Sugar and Company of Korea has now undertaken to produce it for $1.

The Cheil vaccine is of the conventional plasma-derived type and is made under licence from the New York Blood Center. The new process increases by a factor of ten the number of doses obtained for each cubic centimetre of starting plasma, by reducing the number of harsh purification steps and starting with more potent plasma. According to Dr Alfred Prince of the centre, another task force member, the vaccine has been tested in more than 1,000 volunteers for safety and immunogenicity, and long-term effectiveness trials are, at present, under way in several countries.

Demonstration programmes will be mounted in Indonesia, Thailand, Brazil and one African country. Over three years the group expects to vaccinate 1.2 million newborn babies. Vaccinating babies is most effective because those infected before the age of three are far more likely than others to become infectious carriers.

The task force includes members coordinated by a non-profit organization, Program for Appropriate Technology in Health, which will also help governments and international organizations to implement their own programmes.

Costs for the three-year demonstration programme are estimated at $12.5 million; a substantial part has already been promised by charitable foundations, and the health programme is "confident" that the remainder will be found.

Tim Beardsley

US biotechnology

Research centre gets fixed home

Washington

ON an unseasonably cold, blustery day last month in the Washington suburb of Rockville, the proud parents of the Center for Advanced Research in Biotechnology (CARB) held a ceremonial groundbreaking of what will some day be CARB's new home. A joint venture of the National Bureau of Standards, the University of Maryland and the local county government. CARB will focus initially on protein engineering and rational drug design.

For now, CARB is housed in temporary quarters in Gaithersburg, Maryland, on the campus of the Bureau of Standards. CARB director Kevin Ulmer says the centre will take advantage of the bureau's expertise in fine measurement technologies, essential to making detailed protein structural determinations. The bureau also possesses a CYBER 205 supercomputer needed to make calculations for modelling of atomic interactions in protein molecules. Ulmer says that CARB will provide the bureau with a window into biotechnology, an area in which it has not traditionally had a very large presence.

For the University of Maryland, CARB is only one aspect of a major thrust in the realm of biotechnology. The Maryland Biotechnology Institute includes centres devoted to marine biotechnology, agricultural, medical and ethical aspects of biotechnology in addition to CARB. The university's 1987 budget allots $5.6 million to these institutes, including provision for 75 full-time positions. Rita Colwell, university vice president for academic affairs, says Maryland is a logical place to house these centres, because the 50-mile corridor between Washington and Baltimore is rich in biotechnology companies and government research laboratories.

Montgomery County, where CARB is located, is also trying to take advantage of the population characteristics of the Baltimore–Washington corridor. The county has created the Shady Grove Life Sciences Center, and is providing funds and space to both the University of Maryland (50 acres) and John Hopkins University (35 acres) for new campuses, has begun work on what will be a teaching facility for high-technology graduate programmes, and the University of Maryland has begun CARB. County officials say the idea is to compete with the San Francisco Bay area and the Boston area for biotechnology companies. Montgomery County is home for the National Institutes of Health, the National Bureau of Standards and the Food and Drug Administration. Some four dozen biotechnology companies have already located in the county.

CARB's major initial thrust — protein structure — has in just the past two years been an area of intense interest. Thomas Steitz of Yale University calls the resurgance of structural biology "spectacular". Steitz attributes the renewed interest to strides in molecular biology that have provided large quantities of molecules of interest for study. Ability to perform site-directed mutagenesis has spurred efforts to derive a theoretical understanding of how DNA sequence changes would affect the structure of the resulting protein.

But the increased interest in structural biology means that CARB will face stiff competition for personnel and funds. CARB has offered positions to seven scientists to form its senior staff, but so far only one has accepted. Ulmer admits it is hard to persuade industry to enter collaborative projects without a full staff on board. The National Bureau of Standards and the University of Maryland will together provide CARB with about $3 million annually, but Ulmer hopes to increase that figure several-fold through contracts and grants from both government and industry.

Raymond Salemme, who directs Du Pont's efforts in protein structure, says

Information technology

New institute

FINAL preparations are being made for the opening of Britain's first higher education institute devoted to the research and teaching of information technology (IT), the marriage of telecommunication and computer disciplines, supported entirely by industry and by its users.

The new institute, to be based in Milton Keynes in Buckinghamshire just north of London, has been possible through an initial tranche of £3.5 million from more than 25 companies in the IT field, in an attempt to reduce the growing shortage of IT skills in the United Kingdom.

The institute has been created in partnership with the Cranfield Institute of Technology in Bedfordshire, whose vice-chancellor, Sir Henry Chilver, has been the chief driving force behind the project.

The institute is a new dimension in British higher education and one that has provoked detectable hostility from some in established academic institutions who suspect that industrially-funded education, the preference of the Prime Minister Mrs Margaret Thatcher, will undermine the already shaky finances of the British educational establishment.

Supporters of the institute, however, consider that its curriculum will be more appropriate to the needs of an industry already desperately short of key skills, and that its existence will lighten the burden now falling on traditional education institutions, which have been unable to satisfy the appetite of industry.

The future of the institute will depend on whether it proves to be commercially viable, which is reflected in the management structure. A managing director, Dr Allan J. Fox, reporting to a board, replaces the conventional hierarchy of vice-chancellor and senate. Fox was previously deputy director (applied physics) at the Royal Signals and Radar Establishment at Malvern.

Initial areas of research are to include software engineering, parallel processing, artificial intelligence in computers and the communications of computers linked through networks. **Bill Johnstone**

Alaskan environment

Long row ahead

Washington

THE first salvo in what is likely to be the major energy debate of 1987 was fired last week when the Department of the Interior recommended leasing 1.5 million acres of the Arctic National Wildlife Refuge in Alaska for oil and gas exploration. The government believes the land lying along the northern Alaska coast could be "the most outstanding oil and gas frontier in North America". But environmental groups call the same land a unique treasure house of wildlife, and doubt whether the area can be developed without severe ecological disruption.

Since 1980, when Congress set aside 17.9 million acres of Alaskan land for a wildlife refuge, the government has been struggling over what to do about its coastal plain. Environmental groups wanted the land declared a national wilderness, putting it off limits for oil and gas exploration. But the Interior Department's report proposes a full leasing plan, arguing that industry can conduct its operation in a way that will minimize environmental damage.

The report argues that the national need for oil and gas makes development of the coastal plain essential. A 150-mile pipeline would carry oil from the eastern boundary of the coastal plain to the head of the trans-Alaska pipeline at Prudhoe Bay.

Alaskan politicians are enthusiastic about the Interior Department's plans. The Alaskan economy is in the doldrums. The government's report maintains that oil exploration in the coastal plain will reduce US dependence on foreign oil, enhance national security and bring a more favourable balance of trade.

The coastal plain is inhabited by a wide variety of fauna, including bears, polar bears, musk oxen, foxes, sheep and migratory birds. But development will have the greatest impact on the Porcupine herd of some 180,000 migratory caribou. Appromost industries have been disinclined to contract with start-up operations, choosing instead to launch in-house divisions. Salemme believes the interest now being shown by the Howard Hughes Medical Institute in structural biology will create academic laboratories that will compete with CARB.

Interest in protein structure is not limited to the United States. Japan has created the Protein Engineering Research Institute that Ulmer says will have an annual budget of $10 million. In Europe, John Tooze of the European Molecular Biology Organization says that more attention will be directed towards structural work at his facility. **Joseph Palca**

---

on oil leases

ximately 272,000 acres, or 78 per cent, of the herd's core calving area lies within the section of the refuge proposed for leasing. The report acknowledges that the impact on caribou number is "uncertain" even if care is taken to avoid disrupting the herd's movements. The report cites successful industry wildlife management at nearby Prudhoe Bay as evidence that the caribou population will not be decimated.

But Susan Alexander of the Wilderness Society says the experience at Prudhoe Bay is not relevant to the coastal plain. She argues that the principal herd at Prudhoe Bay, the Central Arctic herd, is resident, whereas the Porcupine herd is migratory. The Porcupine herd is also much larger than the Central herd but has a relatively small core calving area. Both factors, Alexander says, make the Porcupine much more susceptible to environmental change. Although assistant Interior secretary William Horn has proposed leasing the calving area last, Alexander says the damage will already be done if leasing is permitted on the range.

There is a 60-day public comment period for the report that expires on 23 January, after which it will go to Congress, sometime in the spring. Congress must approve the plan, and will probably hold its own hearings.

Representative Morris Udall (Democrat, Arizona) is likely to reintroduce legislation that would convert the land from wildlife refuge to wilderness, forestalling energy exploration. Similar legislation died earlier this year. As chairman of the Interior committee, Udall will have a powerful voice in the ultimate disposition of the land.

Alexander says a "suck-America-dry-first" policy is folly, because even Alaskan oil reserves will run out eventually. The coastal plain is a unique place, she says, one "we can afford to leave until the end". **Joseph Palca**

Laboratory fraud

Another damned by publications

San Diego

THE University of California is still trying to pick up the pieces after what appears to have been an extensive scientific fraud which has left the medical school at the campus here (UCSD) severely shaken. Although the case has nothing in common with that which came to light two weeks ago at the Dana-Farber Cancer Institute of the Harvard Medical School (see *Nature* **324**, 197; 1986), scientists are asking if the two cases may have been caused by pressure on investigators to publish.

The now-published report of the university's investigating committee, chaired by thoracic surgeon Richard Peters, into the San Diego fraud has revealed that as many as 68 papers published by former UCSD cardiologist Robert A. Slutsky must be considered questionable. Some of them contain fabricated data.

One of the distinctive features of the Slutsky case is Slutsky's use of what the committee calls "gift authorships", the practice of listing as co-authors colleagues who had made little or no contribution to the work, who could not vouch for its accuracy and who, in some cases, were not aware that their names had been used until the papers carrying them had been submitted for publication.

According to the committee, in the period 1983–84 Slutsky (then 37) was averaging one published paper every 10 days. The committee notes that this pace of publication went unremarked.

Although some of the co-authors were at the outset unaware of the 'gift' of co-authorship, the committee says that others questioned during its year-long investigation freely admitted to the practice of listing as co-authors senior scientists who had contributed nothing more than equipment or laboratory space to the work reported. This practice, it says, is "commonplace and condoned".

As a remedy, the committee recommends that each university department should take responsibility for specifiying the participation of each faculty author in the work described in every paper submitted for publication. It says that co-authors should be responsible for reviewing all manuscripts carrying their names and should be able to vouch for the accuracy of the work. During the investigation, the committee says, it summoned Slutsky's co-authors to defend the papers on which their names appeared, but some of those concerned objected. Peters says that the notion of "innocent until proven guilty" does not apply in science.

Peters also says that the problems at the San Diego medical school could and should have been discovered earlier. The failure, he says, is less in the research enterprise than in the senior scientists who lead it. Paul Friedman, the medical school's dean for academic affairs, said that some of Slutsky's colleagues had been suspicious of the quality of his work but that they "were simply bowled over by his energy, his enthusiasm and the size of his bibliography".

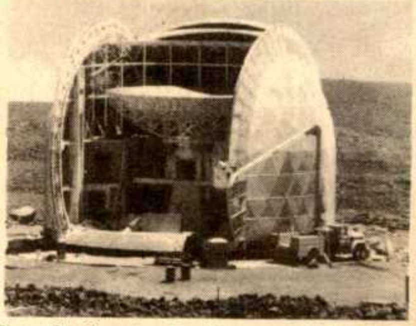
Slutsky's fabrications were apparently uncovered when he was being considered for possible promotion. Radiologist Elliot Lasser set two Slutsky papers side by side and noticed that, in two overlapping data tables dealing with control animals, the numbers of animals were different but the standard deviations were identical.

In retrospect, it has emerged that Slutsky's earlier work shows small discrepancies, as when the same data have been reused or the statistical significance of a result exaggerated. In the end, says Friedman, it emerged that the experiments simply had not been done.

Friedman says that Slutsky was caught because he did not understand statistics, and that he had survived for so long only because journals' reviewers also lack statistical expertise. Slutsky seems to have been successful because he stuck to experiments that he knew would work.

But Slutsky regularly ignored requests for supporting data, for which reason Friedman suggests that the retention of raw data should be a university requirement. He also says that young researchers should be given an estimate of what would be a realistic output of publications.

Friedman is concerned that Slutsky's unsound papers may be only the tip of an iceberg. He says that only the blatant fabrications are usually caught, so that nobody knows how many more cautious investigators cook their data in ways that escape detection. He says that the familiar use of expressions such as "dry-labbing" suggests that fraud is "not an extremely rare phenomenon" **Robert Locke**



THE California Institute of Technology (Caltech) Submillimeter Observatory (shown during construction), a 10.4 metre dish that will operate at wavelengths down to 200 micrometres, was dedicated on 22 November on Mauna Kea, Hawaii. ☐

Science teachers

British attempt solutions

The shortage of teachers of science, technology and mathematics in British schools prompted last week a string of proposed remedies, including the appointment of retired engineers and scientists to teaching posts, dropping the entrance requirements of universities and increasing the pool of supply teachers.

These proposals emerged at a conference in London, where representatives from 51 universities, polytechnics and colleges responded to the British government's consultation document on the shortage of specialized teachers. The government paper, *Action on Teacher Supply in Mathematics, Physics and Technology*, published in August, took the gloomy view that the present crisis will deepen as a technologically based society demands more from its workforce.

Britain needs more than 2,500 physics teachers alone to make up the shortfall, claims the Association of Science Education. At least 20 per cent of physics teaching is done by the unqualified. It is the same in mathematics and technology, where nearly a third of mathematics and a half of technology teachers are inadequately trained.

A statement signed by the universities, polytechnics and colleges says that "shortages of well qualified mathematicians, physicists, computer specialists, engineers and technologists are acting as a brake on economic growth, pushing up pay above the level of inflation and sucking the best staff out of teaching jobs in schools and higher education." The situation is aggravated by very few students applying for courses in these subjects.

The University Grants Committee, which controls the purse strings of the British universities, suggested that universities should be able to lower their admission requirements for courses for teachers in mathematics and physics; £1 million will be available in 1987–88 "for increasing the supply of mathematics, physics and CDT teachers in schools".

The Treasury announced last week that companies who second professionally qualified employees to assist education will qualify for tax relief. Industrially based mathematicians, physicists and engineers near retirement age should be encouraged to retire early and switch to teaching, says the council. **Bill Johnstone**

- British teaching problems are but a reflection of those in the United States. Last week, 90 of 126 invited research universities joined the Holmes Group; while 25 rejected the invitation. The group calls for more graduate teachers and stronger links between schools and universities.

Biotechnology companies

One falters, another forms

Despite its recent successful sale of stock worth $40 million, Biogen, one of the first and best-known biotechnology companies, is not sitting pretty. Indeed, it has decided that it can no longer keep the Geneva end of the business going. Bryan Sautelle-Smith, managing director of Biogen SA in Geneva says that he is in discussion with a number of companies that might become joint partners in his company. The alternative is that Biogen SA is sold.

The financial results for the third quarter of 1986 illustrate the problem faced by Biogen NV, the parent company of both Biogen SA and Biogen Inc. in Cambridge, Massachusetts. Biogen NV has already accumulated a net loss of $20.5 million this year, having lost almost as much in the whole of 1985. Worse still, 1986 revenue so far amounts to only $8.5 million, half the figure of this time last year. Jim Vincent, who replaced Walter Gilbert as chief executive of Biogen just over a year ago and who is now also chairman of the company, says that the decline is mainly due to the scheduled end of payments from collaborators and to fluctuations in revenues from licensing agreements. Vincent is determined to break even as soon as possible and with few prospects of greatly increasing income in the immediate future his only option is drastically to cut costs.

Like most biotechnology companies, Biogen is faced with the fact that its revenues from sales are still small. It receives some income from the sale of hepatitis B core antigens for diagnostic test kits. And it is expecting to receive increasing royalty income from the sale of alpha interferon by the Schering Corporation, for which it developed the product. (Schering, like Monsanto, has recently sold its substantial shareholding in Biogen.) Alpha interferon received UK approval for use in hairy-cell leukaemia this year and is approved in some countries for use in the treatment of multiple myeloma and Kaposi's sarcoma, a condition often found in patients suffering from AIDS (acquired immune deficiency syndrome).

Hopes for increased sales revenue hinge in particular on the use of gamma interferon in the treatment of rheumatoid arthritis. Trials in West Germany are claimed to show its therapeutic value. Bioferon, a company jointly owned by Biogen SA and Rentschler Arzneimittel GmbH, is hoping soon to receive approval for the sale of gamma interferon for use in rheumatoid arthritis in West Germany where it has patent protection.

While Biogen savours the problems of adolescence, a new biotechnology company is launched this week in Britain. Rising like a phoenix from the ashes left behind when Monsanto decided to axe the UK laboratories of its subsidiary G.D. Searle almost a year ago, British Biotechnology Limited starts with £2.5 million venture capital, about twenty ex-Searle staff and more than a touch of patriotism.

Fundamental to the hopes of the new company, situated in Cowley on the outskirts of Oxford, is the expertise in gene synthesis built up at Searle. Most of modern biotechnology is built on the isolation of genes and their manipulation, but British Bio-technology will synthesize and assemble parts of genes from scratch, enabling the production of hybrid proteins designed to have improved therapeutic properties. Dr Brian Richards, chairman of the new company, says that he has been careful to ensure that the patents for specific aspects of gene synthesis held by Searle do not preclude the pursuit of the general approach. A more enterprising plan is to synthesize a variety of carbohydrate groups that can be attached to the proteins, again in the hope of improving their therapeutic value.

Collaboration with established pharmaceutical companies is the initial goal of British Bio-technology. Two major projects are under discussion. One is for a novel thrombolytic agent, produced via a synthetic gene. The other is for an enzyme inhibitor made by organic synthesis, the other general technique on which the company is founded. With the help of an advisory board that includes leading British experts on AIDS, the company will examine the design of drugs against the disease and antiviral agents in general. Arthritis and bone disease are two other major targets.

The initial £2.5 million, which buys 45 per cent of the company's shares, will support the growth to about 80 staff in the next 18 months or so. Apart from the additional income from contract research, British Bio-technology also expects considerable income from the sale of their synthetic genes and carbohydrates to customers in industrial and academic laboratories.

Both Richards and Dr Keith McCullagh, the company's chief executive, are keen to tap British expertise, not least around the corner at the University of Oxford, and are proud of the 'British' in the company's title. To use it, they had to satisfy the government-run Companies Regulation Office that they were both British and pre-eminent. They will now be flying the flag alongside their neighbours in Cowley, British Leyland, the ailing remnant of the UK car industry.

Peter Newmark

Eastern Europe

Widespread shortage of energy

THE Soviet Union and its European Comecon allies face considerable problems with energy supplies this winter. This is, in part, due to the loss of generating capacity following the Chernobyl disaster. The damaged reactor No.4 is now permanently 'entombed' in concrete, reactor No.3 is still out of service and, throughout the Soviet Union, RBMK reactors have had to be temporarily withdrawn from service to make them 'operator-proof', a process which, according to Dr Valerii Legasov, will not be completed until 1987.

But Soviet hydroelectric generating capacity is also down because of a water shortage in Central Asia. Furthermore, there are shortfalls in oil, gas and coal production, due to lack of maintenance and repair work.

Last month, it was reported at a meeting of the Supreme Soviet's Commission on the Fuel and Energy Complex that the "energy shortage" due to Chernobyl and the lack of hydroelectric power would be "met" by increased efforts from the thermal generating sector. Just what "met" means in this context is not entirely clear, as Yurii Semenov, deputy chairman of the Fuel and Energy Complex Bureau of USSR Council of Ministers, predicted that electricity supplies would be "particularly difficult" in Ukraine, Central Asia and the Transcaucasus.

There are already difficulties with heat supplies in Moscow and Tashkent. Comecon countries that import electricity from the Soviet Union have had their quotas cut. In the case of Hungary, which normally imports almost 30 per cent of its electricity from the Soviet Union, instead of the hoped for 10,500 million kWh this year, there will be a cut of 500 million kWh, of which 280 million will come in the winter months. To replace this, the Soviet Union has offered 168 million m^3 of natural gas, but the Hungarians are not entirely happy about the substitution. At the Comecon summit in Bucharest last month, the Hungarian Prime Minister, Gyorgy Lazar, spoke critically about "occasional disorders in the unified system" of energy supply "in the recent period" — a remark widely interpreted as referring to shortfall in natural gas supplies either due to production difficulties in Siberia, or to supplies earmarked for the Soviet bloc being diverted to Western Europe to earn hard currency.

Restriction of supplies from the Soviet Union has led to some squabbling among the Comecon partners. In Bucharest, Lazar accused unspecified "partners" of the "unsystematic acquisition" of electricity, to the danger of Hungarian supplies. This statement was taken by the Hungarians as a reference to the Romanians, who, the Hungarians say, in early 1985 almost blacked out the entire "integrated Comecon grid" (Eastern Europe plus the Soviet Union west of the Urals) by attempting to draw more electricity than their entitlement.

Romania has recently come under fire from several East European countries on energy issues. Last year, according to the Yugoslavs, the Romanians drew too much generating capacity from the joint Yugoslav–Romanian Iron Gates hydroelectric station. More recently, the Czechoslovak party daily *Rude Pravo* accused the Romanians of failing to supply power station parts which, the paper said, will cause delays in refitting Czechoslovak power stations and cause electricity shortages well into 1987. Despite draconian restrictions on domestic use (only one bulb, maximum 40 W per room, and no heating of rooms or water in the morning or the evening) there are still regular lengthy power-cuts in Romania.

Other Comecon countries have also announced rationing. In Bulgaria, there is a limit on the number of kilowatt-hours allowed per household per month, and a new crime, "theft of electricity", has been defined. In Poland, a major coal-producer, there is said to be "chaos" at the coal shops, and the radio carries daily announcements as to which enterprises may expect energy supplies that day.

The preferred solution to the energy problems of the Comecon bloc is nuclear energy. Chernobyl has not changed this; the only difference is that there has been a major effort by the nuclear authorities in the member countries to assure the public that the VVER reactors they use are of different type from the RBMKs used at Chernobyl, and that, in any case, additional safety measures are being taken.

A new programme of nuclear energy cooperation in the Comecon countries has now been completed, and is awaiting ratification. This will include the construction of nuclear power stations and nuclear district heating stations for cities. Twelve such heating stations are now under construction in the Soviet Union. **Vera Rich**

US defence

Biological war still a threat

Washington

A PENTAGON report* released earlier this month on the threat of biological war from the Soviet Union reflects the growing interest in the subject in US military circles. The report reiterates claims that the Soviet Union is violating the terms of the Biological and Toxin Weapons Convention of 1972. It also provides the most comprehensive evidence so far released supporting the US contention that the release of anthrax in Sverdlovsk in 1979 was related to weapons research.

Without explaining how the evidence was gathered, the report says that as much as 10 kg of dry anthrax spores was released into the atmosphere in Sverdlovsk. The report says that the scope of the clean-up, the number of victims and the involvement of the military conflict with the Soviet explanation that contaminated meat was the source of the outbreak.

The report follows the regular five-year review of the 1972 Convention that took place in September in Geneva. Pentagon spokesman Major Randy Morger says the current report has been in preparation for some time, and had been intended to coincide with the Geneva meeting but was delayed at the printing stage.

According to Brad Roberts of the Center for Strategic and International Studies, the renewed interest in biological weapons issues can be traced to a presidential commission on chemical warfare chaired by former under secretary of state Walter Stoessel. The commission said last year that the United States could not afford to ignore Soviet activities on biological weapons, and should embark on comprehensive defensive research on biological agents. This strategy, the Stoessel commission argued, would deter the use of biological arms.

Although the Pentagon report concludes that the Soviet Union is developing biological weapons, it admits that violations of the 1972 Convention are difficult to identify. Research done for medical, biological and public health reasons is "also relevant to developing disease agents for warfare purposes". But the Sverdlovsk incident, the report claims, indicates activities that cannot be justified as prophylactic, protective or for peaceful purposes as permitted by the convention.

The report offers no further data on the yellow rain controversy, explaining in a footnote that its purpose is to focus on agents of disease rather than toxin weapons. **Joseph Palca**

---

* *Soviet Biological Warfare Threat, Defense Intelligence Agency DST-1610F-057-86, 1986.*

Nuclear risk assessment

Sir—I write about the letter of S. Islam and K. Lindgren[1]. Exposure to nuclear accidents and the question whether nuclear engineering is safe or not, and to what degree, is not a question to be answered by sterile theory. The life, health and general well-being of many depend on the availability of cheap and comparatively safe energy. If silly calculations exaggerate the problems caused by mass hysteria, this is not in the best interest of an advanced technological society.

The conclusions reached by Islam and Lindgren are not only wrong but, worse, misleading for those such as the Greens in West Germany. Here are some obvious flaws.

(1) The authors write that "risk assessments which use data from operating experience are a must". I wish they had followed their own advice. What did they do? They inserted the operating experience (reactor years) of a very mixed sample and two accidents (Three Mile Island and Chernobyl) into some formulae. On such a basis, a hypothetical assessor of earthquake catastrophes would have entered an observation period of 10 years at the time when Shakespeare wrote *King Henry VI*, ten to eleven years after the "London Earthquake", Isaac Newton would have used 107 years when his *Philosophiae naturalis principia mathematica* was published and R.A. Fisher would have taken 470 years in 1950. One may repeat this *Gedankenexperiment* in relation to the catastrophic earthquakes of 1811–12 south of St Louis, Missouri. The authors should also ask themselves what their exposure calculations would have produced had the London or the Missouri earthquakes not occurred.

Backed by decades of risk assessment and using the very extensive data of the two largest reinsurance companies, I can assure the authors that a realistic probabilistic evaluation of the catastrophe potential is possible only if one has plenty of information on all accidents, from the smallest to the very largest; only then can reliable probability distributions be inferred.

(2) The authors have (generously) not counted the Windscale accident of 1957 "because it was a military reactor". What matters is the type and the parameters of a reactor, not whether it is a private or a military reactor.

(3) It is obviously unknown to the authors how large the difference is between Soviet RMBK-1000 reactors, light-water moderated reactors and other types. Accident probability depends not only on such basic differences but on engineering details, quality control and on hazards such as fire, explosion, earthquake, flood and so on, and on human parameters. No greengrocer would conclude from an observational sample of thousands of banana and strawberry-years that he must also sell his coconuts and almonds within a day or so to reduce loss from deterioration. Calculations that lump all types of reactors into one sample are not worth taking seriously.

Risk assessment segregates the total exposure according to parameters controlling accidents, when it is possible to correlate actual damage experience with individual parameters to find out which of them appears to be important, whether the list is complete and how much information on accidents is available. More advanced methods (*cf.* ref. 2) but even comparatively simple stochastic calculations, will show that a very large number of observations is necessary per subsample to produce reliable results.

Regretfully, I must add that I know of only a few cases where planners, designers, engineers or contractors of large and risky projects have consulted those with the largest collection of data on accidents and damage (such as specialized insurers and reinsurers).

Unfortunately, it is not generally realized that mathematical models, hypothetical reasoning or thinking carefully about accidents, using logically plausible scenarios[3] and subtleties such as wondering whether to use Poisson, Pearson, Weibull or Gumbel III, are no cheap substitute for difficult, time-consuming and costly compilations of information on actual failures. The observed failure frequency of civil engineering structures, for example, is several orders of magnitude greater than the safety to be expected from calculations[4]. How large is the error margin if far more exotic projects are designed according to logical, albeit theoretical scenarios?

Herbert Tiedemann
Stefan-Rotthaler-Str. 5,
D – 8052 Moosburg, FRG

1. Islam, S. & Lindgren, K. *Nature* **322**, 691–92 (1986).
2. Wald, A. *Sequential Analysis* (Wiley, New York, 1947).
3. Wilson, R. *Nature* **322**, 768 (1986).
4. Rybicki, R. *Schäden und Mängel an Baukonstruktionen*, **10** (Werner, Düsseldorf 1972).

Avalanche dowsing

Sir—I am grateful to *Nature* for publishing accounts of the scientific status of the shadowy field of paranormal phenomena, an area most serious scientists deliberately ignore.

The Norwegian Red Cross Mountain Rescue Organization officially advocates dowsing as an effective method for finding victims of avalanches. This voluntary organization has about 5,000 members in the mountains during the Easter skiing season. Dowsing has been taught for more than 10 years, under the pretence that the technique is accepted by mountain rescue organizations in other countries. Dowsing exercises have been part of the avalanche safety course taught to career officers of the Norwegian Army, and dowsing rods are part of the Army equipment for avalanche rescue.

On 5 March 1986, 16 young Norwegian soldiers died in an avalanche at Vassdalen in northern Norway. Dowsing was attempted by an Army officer in an area where seven or eight soldiers were later found dead. He got between 20 and 30 signals, of which one or possibly two pointed to a victim. He attributes the poor result to the presence of water and vegetation under the snow but is convinced the technique would work under other conditions. A central representative of the Red Cross has expressed the opinion that the dowser was not sufficiently familiar with the technique. The two positive finds made by dowsers at accidents in snow can both be associated with external visual clues. The results are thus in line with those presented, for example by Foulkes[1] and Marks[2], that in the absence of such clues the outcome of dowsing is no better than a series of guesses.

Fortunately, dowsing at the Vassdalen catastrophe did not interfere with the regular search activities which took place simultaneously.

Dowsing is not accepted by the Norwegian police and the Norwegian Geotechnical Institute, responsible for scientific studies of avalanches in this country, has taken a firm stand against the technique.

Rolf Manne
Department of Chemistry,
University of Bergen,
N-5000 Bergen,
Norway

1. Foulkes, R.A. *Nature* **299**, 163–168 (1971).
2. Marks, D.F. *Nature* **320**, 119–124 (1986).

Wrong dose

Sir—In your recent article "Occupational radiation: British nuclear workers cleared" (*Nature* **323**, 481; 1986), there was a major error. It was reported that for Sellafield workers, "the average radiation dose for those for whom film badge records are available is given as 124 mSv (12.4 mrem) accumulated until the end of 1983". Which dose is correct? A dose of 124 mSv is equal to 12.4 rem. 12.4 mrem is equivalent to a dose of 124 μSv. There is a large difference between 12.4 mrem and 12.4 rem.

T.L. Kellogg
University of Vermont,
Radiological Safety Office,
004 Rowell Building,
Burlington, Vermont 05405, USA

- The smaller dose was that intended. Editor, *Nature.*

Images in and of the brain

It is beginning to be possible to produce images of the brain in action, but interpretation of the pictures obtained remains uncertain.

DRAMATIC pictures of the brain in action are now being obtained by optical recording using voltage-sensitive dyes and by positron emission tomography. Some of them have been published in *Nature* (Blasdel, G.G. & Salama, G. **321,** 579; 1986 and Fox, P.T. *et al.* **323,** 806; 1986); last week, A. Grinvald *et al.* (*Nature* **324,** 361; 1986) reported an optical method for recording cortical activity without the use of dyes, at least when the surface of the cortex can be made accessible.

Producing these images requires sophisticated computer processing; understanding them needs tutored interpretation, involving some of the same neural mechanisms that are the subject of the images. Problems such as these were discussed at a recent symposium* on images and understanding, which brought together people who generate images, from the arts and computer graphics, together with neuroscientists and artificial intelligence people who seek to understand the mechanisms by which the brain analyses images.

The conundrum of how the brain interprets images of the brain was illustrated by Horace Barlow (Cambridge) and Marcus Raichle (St Louis), but the implications were left unexplored. The advantage of the new cortical imaging techniques is that the activity of many neurons can be monitored at once, in the intact brain, so as to give a much more complete picture of the spatial organization of neuronal responses to particular stimuli than could previously be had by painstaking reconstruction, and, with some guesswork, from single-unit recordings. Microelectrode recording is, of course, not superseded by the glamorous new technology, for it remains the best way of characterizing the responses of individual neurons. Rather, the synoptic methods add a new dimension by making it possible to see all the neurons activated at once by a particular stimulus; they provide functional maps of the brain.

What do such maps reveal? It has long been known that the brain itself contains a map of the external world — in several areas of the cortex, there are detailed point-to-point, though not necessarily proportional, representations of sensory fields or of the body surface. A consistent finding is that the world is represented many times over, both within a single area and in several regions; different aspects of the same information are then processed in parallel channels.

Colin Blakemore (Oxford) suggested that one reason for such maps is to bring common functions to adjacent points for efficient processing of information. This may be true for orientation and ocular dominance (the basis of depth perception) in the striate area of the visual cortex, as Blasdel and Salama's voltage-sensitive dye maps suggest. But in the same region, cells responding to colour and direction of movement are physically intermingled yet functionally quite separate and they project to separate secondary visual areas (Hubel, D.H. & Livingstone, M.S. *Nature* **315,** 325; Shipp, S. & Zeki, S. *Nature* **315,** 322; DeYoe, E.A. & Van Essen, D.C. **317,** 58; 1985). The organization at these relatively low levels in the processing hierarchy is probably a sorting process through which the correct combinations become addressed to multimodal processing centres higher up the system.

All the maps of the world are symbolic, the maps of the brain included. Reading them must be learned, just as people must learn to interpret geographical maps. Because it is still early in the development of the new imaging technology, the symbolic conventions are not well established and the viewer has to work hard to translate grey-scale or false-colour images into patterns of cortical activity. What is more, the published images are static snapshots of activity patterns at different times.

One of the most difficult features of the brain to capture and analyse is the pattern of dynamic change in the activity of and interaction between its constituent neurons. Because animals and the world around them are rarely still, the sensory input to the brain is constantly changing. Even when one looks at a static image, the attention shifts to focus on different aspects — and appreciation of the same image on different occasions is tempered by physiological and psychological states.

Others know of the difficulty. The art historian, Ernst Gombrich, pointed out how difficult it is to represent a sequence of movements pictorially without recourse to symbols or verbal explanation. Pictorial instructions are therefore rarely universally comprehensible. A clear example of how the interpretation of symbols depends on cultural context and hence on learning was provided by Monica Parker, choreologist to the Royal Ballet, London. In the Benesh notation for recording movements, a basic diagram representing a simple position with arms outstretched at shoulder height would look very different if adopted by an Indian classical dancer rather than by a Western dancer. The details that give the position its character will depend on the form of expression inherent in the tradition of the interpreter.

Although many of the visual images we create and enjoy are static, the perception of movement is a basic and important aspect of visual analysis. In primates, the motion-processing channel starts in the retina but directional sensitivity is first displayed by neurons in area V1 of visual cortex and further elaborated in the motion-analysing area, called MT.

Neurons in MT project to association areas including temporal cortex where, in monkeys, David Perrett and colleagues at St Andrews have recorded from many visually driven cells selective for aspects of body movements. Their results begin to give a glimpse of how the brain reconstructs complex images. Each cell responds to limited views of the body and limited sets of movements, for instance, to a body facing and moving towards the viewer or to a particular articulation of a limb. Many do not depend on absolute form, but respond similarly to a monkey, a human being or with a pattern of dots, as long as it is moving correctly. Although these higher-order cells combine information about motion, form, colour and depth, each still responds only to selected features of the object. Specific recognition or perception of the body as a whole must be the product of collaboration between many cells. Where and how this occurs for any image, simple or complex, is one of the unsolved mysteries of neuroscience.

As a highly visual species, how human beings perceive the world is bound to be a challenging subject; most research on sensory processing and perception has focused on the visual system. But it was for an artist, Liliane Lijn, to point out that ultimately our perception of the world is multisensory; visual images are reinforced by sound, smell and touch. How and where in the brain all the parallel channels carrying the various aspects of all the stimuli converge are unanswered questions. Beyond that still are our conceptual capabilities, defined by Richard Gregory (Bristol) as the marshalling of ideas on a long timescale to develop long-term strategies, generalizations and abstractions.

Jennifer Altman

---

*International Symposium on Images and Understanding, Royal Society, London, 6–8 October 1986.

Mathematics

One hundred per cent proof

Ian Stewart

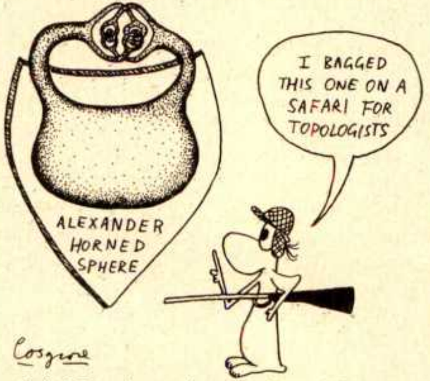
MATHEMATICIANS are a funny lot. They operate in a world of mental abstractions, and many of them appear to have little interest in anything outside it. They pursue abstruse problems on a basis of aesthetic taste, rather than directing their attention at explicit practical goals. Despite this otherworldliness, the results of their deliberations can have enormous practical impact, but that is not my target here, which is the following. To the eternal frustration of their colleagues in neighbouring disciplines, mathematicians are without exception obsessed by the belief that nothing in mathematics is true unless it has been proved, not just beyond reasonable doubt, but beyond unreasonable doubt. Why?

The seamen of ancient Corinth knew that if you tie an overhand knot in a rope, then it cannot be removed except by untying it. Mathematicians remained unconvinced until 1926, when Kurt Reidemeister came up with with an argument that went beyond mere plausibility and established as a mathematical fact the existence of knots that cannot be untied.

Nobody in his right mind would deny that a closed curve in the plane, which does not cross itself, divides the plane into two pieces — the inside and the outside of the curve. Mathematicians were sufficiently sceptical that the question was for many years a well-known and important unsolved problem, stated explicitly by Camille Jordan in 1887, who gave a lengthy, intricate proof. Several other distinguished mathematicians, unhappy with the details, gave their own proofs, but they all contained gaps or mistakes, and the first rigorous proof was given by Oswald Veblen as late as 1905.

In 1852 Francis Guthrie, a student at University College London, wrote to his brother Frederick about a tantalizing little conundrum that he had stumbled across while colouring a map of England. "Can every map drawn on the plane be coloured with four (or fewer) colours so that no two regions having a common border have the same colour?" Experiment rapidly suggests that the answer is yes. But can it be proved? At first this four-colour problem looked fairly easy, the kind of thing that any competent professional ought to be able to polish off with one hand tied behind his back. In 1879 Arthur Kempe, a barrister who belonged to the London Mathematical Society, published a clever argument that seemed to settle the matter. But the problem turned out to be rather more subtle. In 1890 Percy Heawood found a mistake in Kempe's argument, and showed that the best that could be salvaged was a proof that five colours suffice. For almost a century this tiny gap, between five proved and four conjectured, remained unbridgeable. Finally Kenneth Appel and Wolfgang Haken (*Ill. J. Math.* **21**, 429 and 567; 1977) produced their celebrated proof that four colours are always enough. Their method involved testing some two thousand maps for a property called reducibility, which took about 1,000 hours on a fast computer. There are still a few die-hards who do not consider this to be a proof at all, because it uses a computer, and the odd rumour of errors still goes the rounds, but most people are convinced.

My concern here is not the nature of proof. That is a very interesting question, but it would take us too far afield. I also hasten to add that during the initial phase of mathematical creation, rigorous proof is usually left to one side until the basic ideas have crystallized out. My question is: why are mathematicians so hard to satisfy? In other sciences the evidence required to accept a theory is not a logical proof that nothing else could possibly be the case. It is experimental evidence that the predictions of the theory seem to be correct. The stronger the evidence, the happier people are with the theory; and they have no difficulty in operating under the slight possibility that one day the theory will turn out to be false after all. Not seeking absolute truth, they are content with not finding it.


I BAGGED THIS ONE ON A SAFARI FOR TOPOLOGISTS
ALEXANDER HORNED SPHERE
Cosgrove

Not so long ago, mathematics was felt to be exceptional. Its truths, it was said, really are absolute. After a good dose of Gödel, the general opinion changed. Mathematics is only as true as its foundations, and it is impossible to prove that the foundations are 100 per cent solid. Mathematicians have learned to live with that. The natural sciences went through something rather similar but a good deal earlier: eighteenth century fears that scientists knew so much about the Universe that they were doing themselves out of a job also turned out to be unfounded.

But mathematicians still insist on very high standards of logical rigour, and are unimpressed by experimental evidence, even when it appears to be virtually watertight. For example, perhaps the most important open problem in mathematics is the Riemann hypothesis, stated in 1859 by Bernhard Riemann. It is known that the zeta-function $\zeta(z) = 1^{-z}+2^{-z}+3^{-z}+\ldots$ is zero when the complex number z is a negative even number. The hypothesis is that any other zeros have real part equal to $\frac{1}{2}$, that is, they are of the form $z = \frac{1}{2} + iy$. J. van de Lune and Herman te Riele in Amsterdam have just shown, by computation, that the first 1.5×10^9 zeros satisfy this hypothesis. Nobody considers this strong enough to be a proof.

Why do mathematicians take scepticism to such a lunatic degree? There are several reasons. One is that there have been many occasions when intuition or strong experimental evidence has turned out to be misleading. An extension of the Jordan curve theorem states that when a curve divides the plane into two pieces, the inside piece is topologically equivalent to the inside of an ordinary round circle drawn in the plane, and the outside piece is topologically equivalent to the space outside that circle. The natural three-dimensional analogue of this is the hypothesis that when a surface topologically equivalent to a sphere is embedded in space, then the space is divided into two regions, topologically equivalent to the inside and outside of an ordinary round sphere. This is very plausible and natural. Unfortunately, it is not true. The Alexander horned sphere shows the outside space can get very messy indeed.

When a young research student called Stephen Smale first told his supervisor he had proved that you can turn a sphere inside out without introducing creases, he was told to go away and find the mistake. Smale is now one of the world's leading mathematicians, and his theorem is both true and famous.

In the theory of prime numbers the number of primes less than a given number x, denoted $\pi(x)$, is approximated by various analytic formula. One is the so-called logarithmic integral $\mathrm{Li}(x) = \int dx/\log x$. The ratio of $\pi(x)$ to $\mathrm{Li}(x)$ tends to 1 as x tends to infinity, so $\mathrm{Li}(x)$ is a good approximation to $\pi(x)$. What about the error? 'Experimental evidence' shows that $\pi(x)$ is less than $\mathrm{Li}(x)$ for x up to about 10^9 or so. This many successful experiments might be considered conclusive. Mathematicians do not see it that way, and neither did S. Skewes who proved in 1933 that $\pi(x)$ must be *greater* than $\mathrm{Li}(x)$ for some x no bigger than 10^10^10^34, where (for typographical convenience) ^ denotes exponentiation.

This points to a major problem with experimental evidence in mathematics. Evidence based on only a finite number of cases is really pretty thin, and that is usually all the evidence you can get. One billion is a tiny number compared with a googolplex (10^10^100), and most numbers are bigger still. Almost all numbers are immense. The proportion of numbers smaller than Skewes' apparently monstrous number, for example, is zero. Indeed in number theory many estimates on sizes of things grow like, say, log log log *x*. This is 34 (using logs to base 10 for convenience) for Skewes' number. Tiny.

So experience shows that plausible intuitions are not to be trusted, and that numerical evidence is not worth the computer paper it is printed on. The problem is compounded by the way mathematics is constructed. It is like a huge house of cards: theorem piled upon theorem, shored up by lemmas and corollaries, with axioms deep in the foundations acting like concrete pilings and hoped-for conjectures floating up in the clouds, unattached to the main building. If one tiny part is faulty, the whole edifice may collapse. I don't mean this too literally: if a mathematician makes a mistake, it seldom has implications that serious — maybe a balcony falls off. There are internal checks and balances. But to maintain the structural integrity of mathematics it is important not to admit shoddy components.

A standard method of proof, which is extremely powerful and therefore very popular, is 'proof by contradiction'. To prove a statement X the mathematician starts out by denying X, follows up the consequences, and eventually arrives at something that contradicts either itself or some known theorem This shows that the denial of X is false, hence X itself is true. It is rather like offering a gambit in chess in the hope of recouping greater gains later; but, in the words of Godfrey Hardy, the mathematician offers "the whole game". Imagine the dangers of using a result that has not been given a decent proof, however plausible it may be, in the course of such an argument. It does not bear contemplating, and mathematicians therefore try very hard not to contemplate it.

Another reason for wanting rigorous proofs is that their absence is a sign that something important is not understood. The Jordan curve theorem, for instance, led to algebraic topology. This is central to today's mathematics, and is currently producing useful payoff in various scientific areas, notably gauge field theory. On the whole, algebraic topology is not the sort of technique that gets invented in a direct attack on a practical problem. It is not like calculus, say, whose need becomes apparent the moment the real-world problem is formulated in mathematical guise. Obsession with proof is not just pedantry: it is a major source of new ideas and methods.

In science, the experimental method is paramount. It evolved primarily as a way of counteracting the natural human tendency to believe something because one wants it to be true. If scientists allow themselves the luxury of believing a theory because it is pretty, or conforms to their political prejudices, then they cease to do effective science. If an engineer wants to make sure that his bridge will not fall down he does not worry about the aesthetics of the design, or the colour: he checks that the design conforms to acceptable tolerances and he makes sure that the bridge is built to specification. In the same way, a mathematician dare not believe a plausible theorem just because he would like it to be true, and he cannot place great faith in even extensive numerical evidence because there are good reasons to assume that may be misleading. Proof, and proof alone, is the safeguard. ☐

Ian Stewart is in the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Cancer chemotherapy

Progress in understanding multidrug resistance

George R. Stark

RESISTANCE to a wide range of drugs is often found in human tumours that have relapsed after initial regression in response to drug therapy[1]. An understanding of this phenomenon would be important in the search for more effective chemotherapy. On page 485 of this issue[2], Gerlach *et al.* report a striking similarity between the amino-acid sequence of P-glycoprotein, a membrane protein of relative molecular mass 170,000 (170K) whose overaccumulation can lead to multiple drug resistance, and the *hlyB* gene product, a 66K membrane protein that mediates active export of the protein haemolysin from *Escherichia coli.* This homology is sufficiently extensive to suggest functions for the P-glycoprotein. Two other groups have recently reported complete complementary DNAs predicting sequences of P-glycoproteins from mouse[3] and human[4] cells; the results are very similar to those obtained by Gerlach *et al.*[2].

Overaccumulation of P-glycoprotein results from amplification of a small family of closely related genes in multidrug-resistant cells[5,6] (for example, cells resistant to anthracyclines, vinca alkaloids, podophyllotoxins and dactinomycin but not to bleomycin nor most alkylating agents and anti-metabolites). Amplification of many genes occurs at a high frequency in cells of an advanced tumour stage and amplification of oncogenes undoubtedly contributes to malignant progression (see ref. 7 for review). It is likely that a small tumour containing around 10^9 cells already contains a few thousand cells in which the P-glycoprotein genes have been amplified, even before therapy[7]. Worse still, treatment with some anti-tumour drugs or radiation can increase the rate of gene amplification[8] and thus increase the fraction of multidrug-resistant cells.

Overexpression of the P-glycoprotein genes is almost always a result of amplification and is the predominant cause of multidrug resistance in current examples. It is clear that there is a small family of closely related P-glycoprotein genes[6,8] and by screening a complementary DNA library from normal mouse cells, Gros *et al.*[9] obtained a clone for one of the P-glycoproteins that includes the entire coding region. When this sequence was placed in an expression vector and transfected into drug-sensitive hamster cells, selection with adriamycin revealed multidrug-resistant colonies in which only the introduced mouse DNA was amplified. This result shows that overexpression of a single member of the P-glycoprotein gene family confers resistance.

P-glycoproteins are relatively scarce in the membranes of most cells but can be increased by 100-fold or more in multidrug-resistant cells. How can an overproduced membrane protein mediate resistance to a diverse collection of drugs, often unrelated in structure and function? Recent work shows that the drugs accumulate less in resistant cells, probably as a result of an increased rate of efflux rather than a decreased rate of influx, suggesting that the P-glycoproteins may pump many different small hydrophobic molecules out of cells[10,11]. Support for this idea comes from other studies using a derivative of vinblastine that can be photoactivated and appears to bind directly to the P-glycoprotein[12,13]. Note, however, that there are other cases of resistance to single drugs in which, for example, specific pathways for drug influx have been lost[14], and also well-known mechanisms of resistance that have nothing to do with transport[7].

Analysis of the similarity between P-glycoprotein and *E. coli* transport proteins supports and greatly extends the efflux model of multidrug resistance. Sequence analysis of the complementary DNAs reveals that the P-glycoproteins consist of a tandem duplication of a more basic structure[3,4]. Gerlach *et al.*[2] show that the homology with the haemolysin transport protein extends over about 230 amino acids at the carboxy terminus of P-glycoprotein. From work with monoclonal antibodies, it is known that the carboxy terminus is found on the cytoplasmic side of the plasma membrane[15]. By comparison with homologous regions of other transport proteins, it seems likely that part of this cytoplasmic domain functions as a nucleotide-binding site, probably using ATP to provide the energy for active efflux[2-4]. As the sequence of P-glycoprotein consists of a tandem repeat, each molecule would provide two such sites.

In anticipation of a direct study of the function of P-glycoproteins for their role in multidrug resistance, two major mechanisms have been considered[2-4]: (1) direct efflux of hydrophobic drugs through a P-glycoprotein pore, perhaps involving other components of the membrane; or (2) protein-mediated efflux in which the drugs are first bound to a carrier protein which is then expelled from the cell. The latter possibility is suggested by the close homology to the haemolysin transport system, which is known to function in this way. The homology also raises questions about the normal function of P-glycoproteins.

Knowledge of the structure and functions of the P-glycoproteins should lead to new attempts at tackling multidrug resistance in patients. It may be possible to take advantage of the exposure of part of the overproduced P-glycoprotein on the cell surface to target cytotoxic antibody conjugates to these cells after the bulk of the tumour has been destroyed by conventional chemotherapy. As the details of the transport mechanism become better defined, it will be possible to consider the design of specific drugs which can interfere with it; compounds have already been discovered that interfere with efflux of drugs from multidrug-resistant cells[16].

An understanding of the mechanism of multidrug resistance presents the possibility of much greater clinical benefit from the enormous effort that has in the past been put into developing and testing chemotherapeutic agents. □

1. Gerlach, J.H., Kartner, H., Bell, D.R. & Ling, V. *Cancer Surveys* **5,** 25 (1986).
2. Gerlach, J.H. *et al. Nature* **324,** 485 (1986).
3. Gros, P., Croop, J. & Housman, D. *Cell* **47,** 371 (1986).
4. Chen, C. *et al. Cell* **47,** 381 (1986).
5. Riordan, J.R. *et al. Nature* **316,** 817 (1985).
6. Van der Bliek, A.M., Van der Velde-Koerts, T., Ling, V. & Borst, P. *Molec. cell. Biol.* **6,** 1671 (1986).
7. Stark, G.R. *Cancer Surveys* **5,** 1 (1986).
8. Schimke, R.T. *Cancer Res.* **44,** 1735 (1984).
9. Gros, P., Ben-Neriah, Y., Croop, J. & Housman, D.E. *Nature* **323,** 728 (1986).
10. Inaba, M., Kobayashi, H., Sakurai, Y. & Johnson, R.K. *Cancer Res.* **39,** 2200 (1979).
11. Lalande, M.E. Ling, V. & Miller, R.G. *Proc. natn. Acad. Sci. U.S.A.* **78,** 363 (1981).
12. Cornwell, M.M., Safa, A.R., Felsted, R., Gottesman, M.M. & Pastan I. *Proc. natn. Acad. Sci. U.S.A.* **83,** 3847 (1986).
13. Safa, A.R., Glover, C.J., Meyers, M.B., Biedler, J.L. & Felsted, R. *J. biol. Chem.* **261,** 6137 (1986).
14. Fry, D.W. & Jackson, R.C. *Cancer Surveys* **5,** 47 (1986).
15. Kartner, N., Evernden-Porcelle, D., Bradley, G. & Ling V. *Nature* **316,** 820 (1985).
16. Kessel, D. *Cancer Surveys* **5,** 109 (1986).

George R. Stark is at the Imperial Cancer Research Fund, Lincoln's Inn Fields, London, WC2A 3PX, UK.

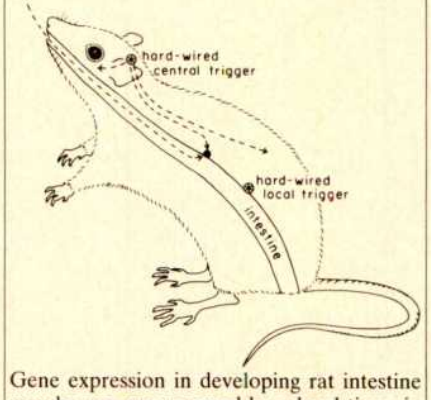
Developmental physiology

Hard-wired local triggering of intestinal enzyme expression

Jared M. Diamond

NORMAL development requires that each gene express itself at the right time and in the right cell. Is proper timing triggered by signals external to the animal, or is it 'hard-wired' within the animal (see figure)? If triggering is hard-wired, is the trigger local, within the tissue where expression will occur, or is it central, for example, in an endocrine gland whose hormones control many peripheral tissues? A recent report by K. Yeh and P. Holt (*Gastroenterology* **90**, 520; 1986) demonstrates the existence of a hard-wired local timer for sucrase expression in rat intestine, which will be a useful model for the identification of timers in other systems.

Infant rats are normally weaned from a diet of maternal milk to solid food around 25 days after birth. Starting around day 18, intestinal activity of lactase (which digests the milk sugar lactose) declines, whereas sucrase (which digests sucrose in solid food) first appears (Henning, S. *A. Rev. Physiol.* **47,** 231; 1985). These changes might be externally triggered by weaning and the resultant change of solutes entering the intestinal lumen: experimentally delaying the age of weaning, for example, is found to delay the decline of lactase.



Gene expression in developing rat intestine may be pre-programmed by a local timer in intestinal cells themselves; pre-programmed by a centrally released timer (for example, a thyroid or pituitary hormone); or not programmed at all but instead triggered by an external event.

The changes might also be under hard-wired central control by hormones that modulate many other developmental changes around the time of weaning. (Pituitary, thyroid and adrenal hormones affect development of intestinal lactase and sucrase, as well as that of many other functions.) Could a hard-wired local timer contribute as well?

Yeh and Holt transplanted isografts of mid-intestine from donor rat pups 0 or 5 days after birth to under the skins of newborn (day 0) host rat pups. Some of the grafts survived and exhibited normal migration rates of differentiating enterocytes along the crypt/villus axis. They found that the control of expression in the isografts differs between lactase and sucrase.

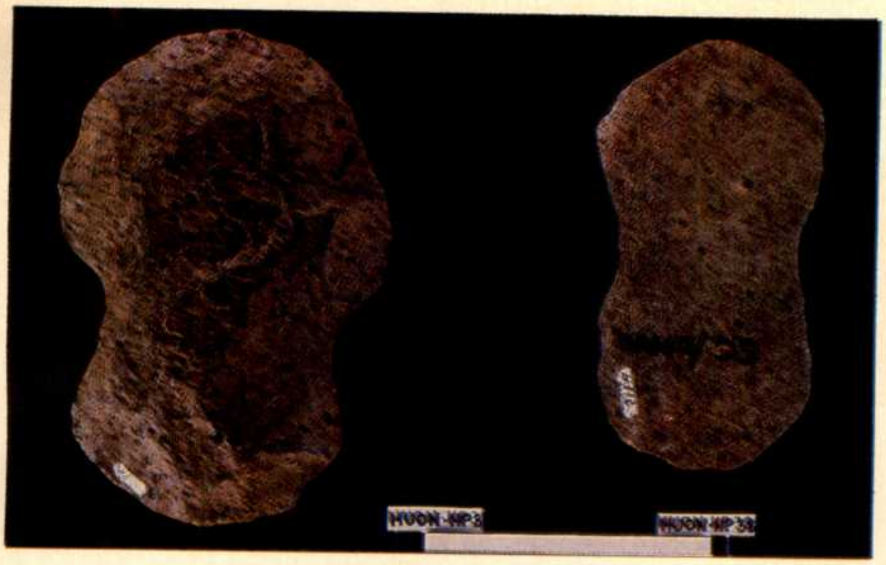
Isograft lactase does not begin to decline until after donor age 28 days (host age 28 or 23 days), although lactase in the host intestine begins to decline as usual around 18 days. Thus, the change in luminal solute input, which was experienced by the host intestine but not by the donor intestinal graft under the host's skin, is an important signal for turning off lactase expression. In agreement with this conclusion, the normal decline in lactase is delayed in bypassed intestinal loops (Tsuboi, K. *et al. Gastroenterology* **80,** 1550; 1981) and is accelerated in rat pups fed an artificial milk formula (Yeh, K. & Holt, P. *Pediatr. Res.* **19,** 963; 1985).

Sucrase activity begins to rise simultaneously around day 18 in donor and host tissue when day 0 donor tissue is transplanted into day 0 hosts. But when day 5 donor tissue is transplanted into a day 0 host, the isograft begins expressing sucrase when the host is only 13 days old and the donor tissue 18 days old. At this time the host tissue expresses no sucrase, which does not appear until 5 days later when the host reaches an age of 18 days.

Thus, donor sucrase expression is not triggered externally by luminal solute inputs (from which the donor intestinal graft was cut off), nor by a hard-wired central timer in the form of host hormones (which would have triggered host and donor expression simultaneously). Nor could the normal absence of sucrase expression at day 13 result from a systemic inhibitor present then, as the graft does express sucrase at a host age of 13 days. Instead, there must be a hard-wired local timer within the donor intestinal tissue that turns on sucrase expression at a tissue age of 18 days. In agreement with this conclusion, sucrase expression begins at the normal time in bypassed intestinal loops and in rat pups from which the pituitary, adrenal or thyroid gland is removed. □

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024, USA.

Discovery of waisted axes in New Guinea



THE recent discovery of waisted axes such as those pictured above indicates that human inhabitants were present in Papua New Guinea more than 40,000 years ago. On page 453 of this issue, Les Groube and his co-workers describe the excavation undertaken (at the south-east end of the Huon Peninsula in New Guinea) and the artefacts discovered there. The archaeological site consists of a series of coral terraces of varying ages, the waisted axes being localized to regions believed to be more than 40,000 years old. This dating was confirmed by thermoluminescence analysis of resident quartz particles. These new data suggest that New Guinea was inhabited at a relatively early stage in the Pleistocene settlement of Australia from the Indonesian–Indochina region, when it and Australia were connected by a land bridge. □

Albert Szent-Györgi (1893 – 1986)

ALBERT Szent-Györgyi, one of the most remarkable biochemists of our century, died in Woods Hole, Massachusetts, on 22 October. The son of a Hungarian well-to-do farmer, he grew up in a comfortable and cultivated household. He had started to study physiology when he was called up for military service in World War I. In the post-war upheavals he wandered from one laboratory to another, learning new techniques from Michaelis and others, until he settled in Holland, first in Leiden and then with H.J. Hamburger in Groningen. He became deeply interested in biological oxidations, especially the catalytic systems involving succinate, fumarate, malate and oxaloacetate. He recognized their wide general significance in metabolism, and his work represented an important preliminary step that helped toward the later demonstration of the total tricarboxylic acid cycle by Hans Krebs.

Szent-Györgyi also became interested in vegetable respiration, and detected a reducing agent, present in many plant juices, which he noted because it caused a slight delay in the progress of oxidations catalysed by peroxidases. That he took notice of this slight delay, which many investigators might have ignored, was characteristic. When Hamburger died, Szent-Györgyi had to find a job elsewhere; by great good fortune Sir Frederick Hopkins, who had followed his work with interest, invited him to Cambridge and obtained a fellowship for him. Here he succeeded in crystallizing the reducing agent from the juice of oranges, lemons and cabbages, and from adrenal glands.

About this time (1932) an enlightened Hungarian minister of education recognized Szent-Gyögyi's ability and invited him back as professor of medical chemistry at the University of Szeged. A young Hungarian, J. Swisbely, who was familiar with the assay for vitamin C, came to work with him. It was soon apparent that the crystalline reducing agent was indeed vitamin C (ascorbic acid). Szent-Györgyi found that Hungarian paprika was extraordinarily rich in this substance, and was able to prepare crystals in kilogram quantities, which he distributed to other researchers throughout the world. It was for his work on ascorbic acid that he received the Nobel Prize in 1937; the citation also noted the importance of his work on the role of dicarboxylic acid anions in biological oxidation systems.

Next he became involved in the study of muscle contraction and the proteins of muscle. On extracting myosin with salt solutions by the standard procedures, he noted that, if the extraction was prolonged, the extract became considerably more viscous, although the total amount of protein extracted was not much greater. This led to the recognition that another protein, later called actin, was associated with myosin. Szent-Györgyi's young co-worker, F.B. Straub, purified actin, while he himself investigated the ATP-induced contractility of actomyosin threads.

Szent-Györgyi carried out this work, which opened great new vistas in muscle biochemistry, during World War II under constant danger from the Nazis, who had become aware of his contacts with the British Secret Service. He was placed under house arrest, and escaped seizure by the Gestapo only by taking refuge in the Swedish Legation in Budapest. When the Nazis broke into the Legation, he fled just in time, and he and his wife went into hiding until the war ended.

When the Soviet armies arrived, Szent-Györgyi and his family were protected and well cared-for on personal orders from Molotov, and later he was brought to Moscow for two months. But he became disillusioned and outraged by the character of the Soviet regime and the atrocities committed by the Soviet armies in Hungary. Although personally well treated by the authorities, he found it impossible to continue the work of his laboratory and so departed for the United States, where he settled at the Marine Biological Laboratory in Woods Hole.

Always ready to strike out into new fields, he became fired with interest in the concepts of quantum mechanics, although he never claimed to have mastered (or even to have actually grasped) the mathematical complexities of the subject. Along these lines, he wrote a small monograph on what he termed 'submolecular biology'.

His major preoccupation during the last decades of his life, however, was with cancer. His views on the origins of cancer differed widely from those of almost all other workers in the field, and government grants for his support, although available in limited amounts, were quite inadequate to support research on the scale that he considered necessary. He continued research with a small group, but his last years must have been embittered with a sense of frustration.

He was always, especially in his earlier years, a man of infinite jest; he loved to announce his ideas and discoveries in a manner that could lead the unwary to believe that he was talking nonsense; but they soon discovered that there was a very solid basis for it all. Like many jesters, he had a deeply tragic vision of the world; he was appalled by the follies of its rulers, teetering constantly on the brink of destruction. Somehow he still managed to cherish hope. Scientist and humanist, he was a unique and unforgettable man.

John T. Edsall

John T. Edsall is in the Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

Evolutionary ecology

Toxins and grazing resistance

Peter D. Moore

ONE of the most useful things about plants, where human beings are concerned, is their capacity to resist the attentions of grazing organisms by the production of deterrent chemicals. These secondary compounds are among the most important controls on grazing insects[1] and by happy chance provide a rich source of drugs, for example, caffeine[2]. Considerable effort is now being directed towards the study of plant/grazer interactions as a potential source of pesticides and drugs, and simply as an area in which co-evolution is most graphically illustrated. Some recent investigations seek to demonstrate the importance of secondary compounds in a range of lower plants, including cycads[3], lichens[4] and algae[5], as well as in higher aquatic plants[6].

Unpalatable meals? Larvae of Eumaeus atala florida *(photograph courtesy of D.C. Lees, London Butterfly House).*

Cycads are well known as plants toxic to vertebrates[7], but the North American cycad *Zamia floridana* is fed upon by the brightly coloured caterpillar of a Lycaenid butterfly, *Eumaeus atala*. It has been supposed that the red and yellow coloration of the larva indicates that it sequesters glycosides for its own defence, as is the case in the monarch butterfly[8], but the study of this particular plant/animal inter-relationship has been hampered by the extreme rarity of the insect. A captive breeding programme has now provided sufficient larvae to permit analysis and Rothschild, Nash and Bell[3] report that both the caterpillars and the adult insects do indeed contain the glycoside cycasin and are therefore toxic to predators. They also emit a pyrazine-like odour which presumably deters enemies. Their conspicuous coloration, also a deterrent, is unusual almong the Lycaenids, suggesting that the association between cycad and butterfly has been a long one. It would be interesting to know whether the cycad receives any benefit from the relationship, perhaps in the form of pollination.

100 years ago

I WISH to suggest a possibility, which if true will considerably alter our views of the facts of chemistry. Let us take the electrolysis of hydrochloric acid.

At present we state the facts thus: Every molecule of hydrochloric acid consists of one atom of chlorine and one atom of hydrogen, the chlorine atom weighing 35.5, the hydrogen atom weighing 1. On passing a current, each molecule is split into these two atoms, each atom carrying a unit charge of electricity.

It is not just possible that we may some day state: A molecule of hydrochloric acid consists of one molecule of hydrogen weighing 1 combined with 35.5 molecules of chlorine each weighing 1. On electrolysis, the chlorine atoms are split from the hydrogen atom, the chlorine atoms each carrying unit charge of electricity, and the hydrogen atom carrying 35.5 charges of electricity. If this is the truth, then all the atoms of the elements are of the same weight, and probably are made of the same "stuff", and we have only two things which condition the properties of the atom — its electrical charge and its electric potential, and Mendelejeff's table becomes a statement of the periodic relationship between these.

From *Nature* **35**, 132: 9 December 1886.

Down the scale of evolutionary complexity, secondary compounds in the lichens are also implicated in herbivore deterrence. Lawrey[4] examined the behaviour of slugs when presented with a selection of lichens and found that they exhibit a distinct preference for certain species — perhaps it would be more accurate to say that they definitely avoid the unpalatable species. Even acetone extracts of the thalli of avoided species are effective in deterring the predators. The fact that it is avoidance of the distasteful rather than preference for the attractive which operates is shown by the higher levels of nitrogen, phosphorus, potassium and calcium in the avoided species. Those lichen species, which should have been most alluring to the slugs, are in fact the most distasteful because of their heavy load of secondary compounds.

In marine algae, there are secondary compounds which could also be involved in grazing resistance. For example, many species of the red alga genus *Laurencia* have sesquiterpenes, compounds which are often involved as insect deterrents in higher plants[9]. There are even parallels to the terrestrial grazers which have an immunity to toxins or even store them for their own protection against predators. The sea hare *Aplysia californica*, a marine mollusc, is though to do this with the sesquiterpenes of *Laurencia pacifica*[5]. But it is evident that not all the grazing effects of herbivores on algae are necessarily harmful to the plant. Zooplankton may stimulate phytoplankton growth rates by local enrichment of water with ammonia[10]; Connor and Quinn[11] propose that growth of some microalgae is stimulated by the mucus of grazing limpets.

But Crews and Selover[5] now suggest that *Erythrocystis saccata*, a red algae semi-parasitic on *L. pacifica*, may take up sesquiterpenes from its host and retain them in the same concentration in its own thallus. They have not demonstrated actual transport, but any other source of this particular suite of compounds seems unlikely. Perhaps the parasite gains some degree of predator immunity from the sequestered defence materials.

It is not only the algae among aquatic plants that are well equipped to deal with the attentions of grazers. Ostrofsky and Zettler[6] show that the same is true of many aquatic higher plants, such as some species of the pondweed genus *Potamogeton*. Possession of deterrent secondary compounds may account for the low level of herbivore attack on such species. *P. crispus* is particularly well supplied with alkaloids and the authors consider that the concentrations present (0.2–0.5 mg per g dry weight) are adequate to account for the general lack of grazing on such aquatic macrophytes. But some terrestrial mammals are known to graze on aquatic plants, for example the moose, which derives most of its sodium requirement from summer grazing in shallow lakes[12]. Perhaps this animal, like the sheep[13], has some degree of resistance to the alkaloids contained in the aquatic macrophytes and thus takes advantage of a relatively underexploited resource. ☐

1. Maugh, T.H. *Science* **216**, 722 – 723 (1982).
2. Nathanson, J.A. *Science* **226**, 184 – 187 (1984).
3. Rothschild, M., Nash, R.J. & Bell, E.A. *Phytochemistry* **25**, 1853 – 1854 (1986).
4. Lawrey, J.D. *Am. J. Bot.* **70**, 1188 – 1194 (1983).
5. Crews, P. & Selover, S.J. *Phytochemistry* **25**, 1847 – 1852 (1986).
6. Ostrofsky, M.L. & Zettler, E.R. *J. Ecol.* **74**, 279 – 287 (1986).
7. DeLuca, P., Moretti, A., Sabato, S. & Gigliano G.S. *Phytochemistry* **19**, 2230 – 2232 (1980).
8. Fink, L.S. & Brower, L.P. *Nature* **291**, 67 – 70 (1981).
9. Harborne, J.B. *Introduction to Ecological Biochemistry* (Academic, London, 1982).
10. Sterner, R.W. *Science* **231**, 605 – 607 (1986).
11. Connor, V.M. & Quinn, J.F. *Science* **225**, 843 – 844 (1984).
12. Botkin, D.B. *et al. Proc. natn. Acad. Sci. U.S.A.* **70**, 2745 – 2747 (1973).
13. Grime, J.P. *et al. J. Ecol.* **56**, 405 – 420 (1968).

Peter D. Moore is in the Biology Department, King's College London (KQC), Campden Hill Road, London W8 7AH, UK.

Galactic evolution

Do globular clusters belong?

Virginia Trimble

Do the globular star clusters really belong to the galaxies they now inhabit, in the sense of having formed in and with them? Yes, according to some of the results presented at a recent symposium*, and no, according to others. The issue arises because these clusters are a bit like dinosaur teeth — small hard things that could have survived a long time under harsh conditions. Thus, they might, alternatively, have formed separately before their host galaxies or as part of other systems that amalgamated to make those galaxies. The issue is important because the globular clusters, unlike dinosaur teeth, are not just very old, but are actually the oldest subjects in the Universe available for study at close range. They ought, therefore, to tell us something of when and how galaxies and stars began.

Globular clusters have so far been seen in about 40 galaxies (spirals, ellipticals and irregulars) and studied closely in half a dozen. The 125-or-so clusters known in our own Milky Way contain, for instance, about 1 per cent of the stars of the spherical halo. Evidently, indicators are needed to show whether the clusters resemble the stars around them more than they resemble each other, and whether their properties (individually or as populations) are correlated with those of their host galaxies. Resemblance and correlation would suggest formation *in situ*, and the opposite would suggest a more complex process.

The evidence, unfortunately, is contradictory and rather evenly balanced. Starting with our own Galaxy, the globular clusters (according to R. Zinn, Yale University) do generally resemble the other halo stars in their distribution of numbers as a function of radius, in their average heavy-element abundance and its gradient with galactocentric distance, and in their dynamical properties (net rotation, velocity dispersions in three dimensions and average orbital eccentricity). Differences exist, however, in the periods of variable stars as a function of their composition (G. Wallerstein, University of Washington) and in the metallicity distribution, in the sense that the field stars have a larger representation at very high and very low values than do the clusters (J.B. Laird, University of North Carolina).

Other indicators that the clusters 'belong' to their hosts are: the likelihood that galaxies more massive than the Milky Way, including M31, NGC5128 and M87, have redder, more metal-rich clusters (F. Fusi Pecci, University of Bologna; H.C. Harris, US Naval Observatory; J. Huchra, Center for Astrophysics, Cambridge, Massachusetts), though the high value for M87 was disputed by J. Nemec (University of British Columbia); the two-standard-deviation result that central (massive, cD) galaxies in clusters have more than their fair share of globulars (W.E. Harris, McMaster University); and the undisputed fact that gas-rich galaxies, including M33 and the Magellanic Clouds, unlike the Milky Way and similar spirals, are still producing massive clusters (though not quite such massive ones as the true, old globulars; E. Olszewski, Steward Observatory; C. Christian, Canada–France–Hawaii Telescope).

In contrast, several other observations suggest that the clusters did not originate with the galaxies they now inhabit. First, the number of clusters at a given luminosity, the function $N(L)$, seems to be much the same, a gaussian with the same width and a peak at $M_v = -7.4$, in nearly all galaxies examined; at any rate, the assumption that this is so leads to a consistent distance scale out at least as far as the Coma cluster (D.A. Hanes and D.G. Whittaker Queen's University, Ontario; L.A. Thompson and F. Valdes, Institute for Astronomy, Hawaii). Second, the clusters in M31 (Fusi Pecci) and M87 (Huchra) do not show the kind of gradient of abundance from centre to periphery seen in the Milky Way. Both items suggest that the clusters did not know where they were going to end up when they formed.

Finally, and most striking, the clusters in some galaxies differ from the non-cluster stars at the same radial position in a number of related ways. The cluster populations are less centrally condensed than the underlying star light, scaling as $r^{-2.5}$ rather than as r^{-3} in M87, M49 and other large elliptical galaxies[1], where r is the distance from the centre of the galaxy concerned. They are more spherically distributed around the galactic centre than the background stars in disk galaxies including NGC3115, 4594 (the Sombrero) and 7814 (Harris). They are bluer than the star light at the same positions both in M87 (J. Cohen, Caltech, confirming earlier work[2]) and in several dwarf ellipticals[3]. And they display a larger velocity dispersion than the other stars at the same radius in NGC5128[4] and M87[5]. The net impression is that the clusters constitute an older, more primitive population, the bluer colours resulting from lower metal abundancies; and the greater extension, sphericity and velocity dispersion representing a lesser degree of collapse and dissipation.

This could mean that the clusters are pregalactic and primordial[6,7], or that they formed as part of larger units that merged to make modern galaxies (R.B. Larson, Yale University). But it could also mean that the globulars formed *in situ* as galactic halos collapsed, their population statistics having been gradually modified as the most vulnerable clusters were torn apart tidally during close approaches to the galactic centres (J.P. Ostriker, Princeton University; M.J. Rees, Institute of Astronomy, Cambridge; D. Chernoff, Cornell University). Tidal disruption would systematically remove central clusters (hence the shallower density profile) and ones on eccentric orbits. The remaining circular orbits then give the impression of increased velocity dispersion at any given off-centre point. To account for the colour difference, the metal-rich clusters must also have been systematically removed, which is also plausible, at least in galaxies with radial abundance gradients. Apparently, then, all the scenarios can still accomodate all the data and no-one suggested unambiguous clues that might be sought in the near future. □

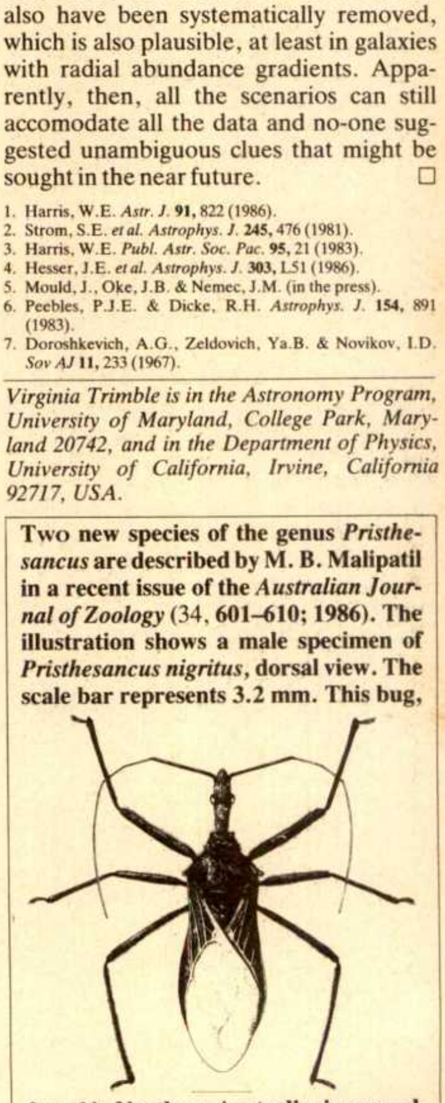
1. Harris, W.E. *Astr. J.* **91,** 822 (1986).
2. Strom, S.E. *et al. Astrophys. J.* **245,** 476 (1981).
3. Harris, W.E. *Publ. Astr. Soc. Pac.* **95,** 21 (1983).
4. Hesser, J.E. *et al. Astrophys. J.* **303,** L51 (1986).
5. Mould, J., Oke, J.B. & Nemec, J.M. (in the press).
6. Peebles, P.J.E. & Dicke, R.H. *Astrophys. J.* **154,** 891 (1983).
7. Doroshkevich, A.G., Zeldovich, Ya.B. & Novikov, I.D. *Sov AJ* **11,** 233 (1967).

Virginia Trimble is in the Astronomy Program, University of Maryland, College Park, Maryland 20742, and in the Department of Physics, University of California, Irvine, California 92717, USA.

---

* *International Astronomical Union Symposium No.126, Globular Clusters.* Harvard University, 25–29 August 1986.

Two new species of the genus *Pristhesancus* are described by M. B. Malipatil in a recent issue of the *Australian Journal of Zoology* (34, 601–610; 1986). The illustration shows a male specimen of *Pristhesancus nigritus*, dorsal view. The scale bar represents 3.2 mm. This bug,



found in Northern Australia, is generally black with two pale distal antennal segments. Like other members of the order Heteroptera, this specimen has two pairs of wings and mouthparts modified for piercing and sucking. Many heteropterous insects are vectors of plant and animal diseases. □

Clinical pharmacology

New hopes for the treatment of coronary heart disease

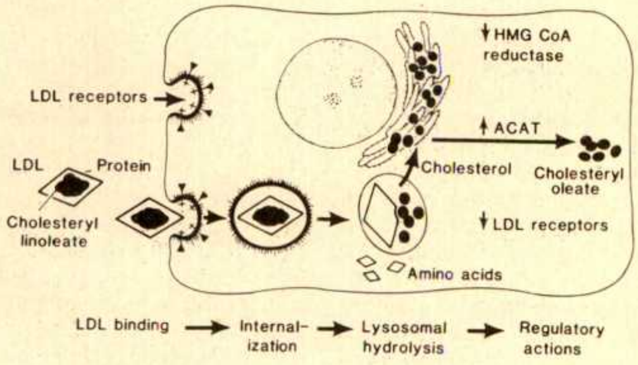
Gilbert Thompson

A RECENT 6-year survey[1] of more than 350,000 middle-aged men in the United States confirms that both total mortality and death from coronary heart disease are strongly correlated with the concentration of cholesterol in the blood. Death rates rise steeply once serum cholesterol exceeds 6.5 millimoles per litre, and values above that level are found in 15 per cent of the population. In the United Kingdom, new epidemiological data[2] indicate that there are twice as many men with this degree of hypercholesterolaemia. In most instances this problem can be managed by dietary means, but a significant minority of individuals will require drug therapy. A new agent that reduces the concentration of cholesterol in the blood is described by A.H. Underwood and colleagues on page 425 of this issue[3].

The suggestion that hypercholesterolaemia causes coronary heart disease is confirmed by the atherogenetic effect of increased levels of low-density lipoprotein (LDL) cholesterol and the protective effect of high levels of high-density lipoprotein (HDL) cholesterol. Drugs that lower LDL and increase HDL are used to treat familial hypercholesterolaemia. People with this condition (0.2 per cent of the population of the United Kingdom) have a propensity to accelerated atherosclerosis and premature death from coronary heart disease. M. Brown and J. Goldstein discovered that this disorder represents a genetically determined defect of specific cell-surface receptors for LDL and have elucidated the structure and function of the receptors.

As illustrated in the figure, the LDL receptor provides a path of entry for exogenous cholesterol into the cell and reduces endogenous cholesterol synthesis through down-regulation of the rate-limiting enzyme β-hydroxy-β-methylglutaryl coenzyme A (HMG CoA) reductase[4]. The number of LDL receptors expressed and the activity of this enzyme in the cell have complementary roles in maintaining cholesterol homeostasis. This process is disturbed in familial hypercholesterolaemia where defective function of the receptors results in diminished uptake from plasma and decreased degradation of LDL, processes which normally take place mainly in the liver. The LDL receptor is a glycoprotein of relative molecular mass 160,000 (160K), the gene for which is on chromosome 19, whereas HMG CoA reductase is a 97K glycoprotein, the gene for which is on chromosome 5 (ref. 5). Several mutations of the gene encoding the receptor have been described[6], any of which can give rise to familial hypercholesterolaemia, but no mutations of the gene encoding HMG CoA reductase have been identified.

Hitherto the ability to stimulate hepatic LDL receptors has depended on the use of insoluble anion-exchange resins such as cholestyramine. These resins bind bile acids within the intestine and thus impair their re-absorption into the blood. To make good this loss, the rate of conversion of cholesterol to bile acids in the liver increases, the consequent need for additional cholesterol being met partly by increased expression of LDL receptors and partly by a compensatory rise in cholesterol synthesis.



A cultured fibroblast showing binding to receptor, endocytosis and lysosomal degradation of LDL, resulting in hydrolysis of protein to amino acids and of cholesteryl ester (linoleate) to free cholesterol. This leads to down-regulation of HMG CoA reductase and LDL receptors. Excess free cholesterol undergoes esterification to cholesteryl oleate by the enzyme acylcholesterylacyltransferase (ACAT). (Reproduced from ref. 11 with permission from M.S. Brown and J.L. Goldstein).

The LDL-lowering effects of cholestyramine can be markedly enhanced by concomitant administration of a new class of drugs that block cholesterol synthesis by competitively inhibiting HMG CoA reductase. Like cholestyramine, these compounds promote receptor-mediated LDL catabolism but are much easier to take. They include mevinolin (lovastatin) and synvinolin, and look set to revolutionize the treatment of heterozygous familial hypercholesterolaemia[7,8]. Unfortunately homozygotes, with fewer functional LDL receptors, respond less well.

The new work of Underwood *et al.*[3] in this issue is based on the well-known fact that thyroid hormones lower serum cholesterol. In attempts to use this effect therapeutically, avoiding undesirable side-effects by administering D-thyroxine instead of L-thyroxine, even D-thyroxine proved dangerous, causing cardiac arrythmias when given to patients with coronary heart disease, and its use has been largely discontinued. Underwood *et al.*[3] have discovered a thyromimetic agent, SK&F L-94901, which has significant cholesterol-lowering properties in animals. Their data suggest that this compound is equipotent to triiodothyranine in its effect on the liver but one thousand-fold less active in its effect on the heart. Although its development is still at an early stage, the outlook seems promising. The mechanism of the cholesterol-lowering effect presumably involves enhancement of receptor-mediated LDL catabolism, which also occurs during treatment of hypothyroidism with L-thyroxine[9]. The possibility that it stimulates selectively hepatic LDL receptors is intriguing but if, like thyroid hormones, SK&F L-94901 also increases HMG CoA reductase activity[10], it might be necessary to administer it with an HMG CoA reductase inhibitor. Such a combination could be just what is needed if diet-resistant hypercholesterolaemia is to be treated as routinely and effectively as hypertension, as the evidence[1] suggests it should. ☐

1. Martin, M.J. *et al. Lancet* **ii**, 933 (1986).
2. Shaper, A.G. & Pocock, S.J. *Br. med. J.* **291**, 480 (1985).
3. Underwood, A.H. *et al. Nature* **324**, 425 (1986).
4. Brown, M.S. & Goldstein, J.L. *Sci. Am.* **251**, 52 (1984).
5. Lindgren, V., Luskey, K.L., Russell, D.W. & Francke, U. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8567 (1985).
6. Tolleshaug, H., Hobgood, K.K., Brown, M.K. & Goldstein, J.L. *Cell* **32**, 941 (1983).
7. Thompson, G.R., Ford, J., Jenkinson, M. & Trayner, I. *Q. J. Med.* **60**, 801 (1986).
8. Mol, M.J.T. *et al. Lancet* **ii**, 936 (1986).
9. Thompson, G.R. *et al. Proc. natn. Acad. Sci. U.S.A.* 78 2591 (1981).
10. Myant, N.B. *The Biology of Cholesterol and Related Sterols* (Heinemann, London, 1981).
11. Brown, M.S. & Goldstein, J.L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3330 (1979).

Gilbert Thompson is in the MRC Lipoprotein Team, Hammersmith Hospital, Du Cane Road, London W12 0HS, UK.

Astrophysics

Big, repulsive boson stars

Robert Walgate

THERE is normally little need to remark that all familiar stars — even quark stars and neutron stars — are made of fermions, particles of half-integral spin whose quantum statistics forbid them to enter the same quantum state. The result of this well-known fact is the 'degeneracy pressure' which keeps the fermions in different states like the electrons in an atom, and prevents even quite massive, relativistic stars from collapsing under gravity. But is it possible to conceive of a stable gravitating star made of bosons, the integral-spin particles which can form a single quantum state of, in principle, unlimited intensity? The theoretical Universe is, after all, beginning to be filled with exotic particles of many kinds from the bosonic 'superpartners' of the quarks and leptons to Higgs bosons and 'ghost' matter; it is conceivable that if boson stars exist then some of the supposed missing matter of the Universe could be in the form of such objects. In fact, as M. Colpi, S.L. Shapiro and I. Wasserman report (*Phys. Rev. Lett.* **57,** 2485; 1986), the answer seems to be yes.

Initial calculations supported by refined work with general relativity led R. Ruffini and S. Bonazzola (*Phys. Rev.* **187,** 1767; 1969) to suggest that boson stars would be possible. But by astronomical standards these early calculations gave the stars a tiny mass, of the order of the square of the Planck mass divided by the boson mass (around 300 million tons, or 10^{-19} solar masses for 1 GeV bosons), and a radius of the order of the Compton wavelength of its constituent bosons (1 fm or so for 1 GeV). It was difficult then to argue for a real astronomical role for such objects. The new work of M. Colpi *et al.* shows that the earlier conclusions — which assumed non-interacting bosons — can be drastically modified if the bosons have the slightest repulsive force among them.

The new work thus demonstrates that boson stars of in principle any mass might be assembled, depending on the interaction strength and mass of the constituent bosons. For example, it follows that a unit interaction strength (of the same order as the strong interaction) and a 1 GeV boson mass would lead to a boson star not of 1 fm, but of 100 m radius, and to the corresponding relativistic maximum mass of 0.1 solar masses. If the bosons were the putative axions of 10^{-5} eV mass, the star would be a crazy 10^{28} times bigger; whereas smaller interaction strengths could produce an object of, for example, the mass of Jupiter. This range indicates at least that boson stars must now be considered candidates to join the observers' menu of theoretical exotic stars.

Why is the interaction important? Essentially because any repulsive force of the order of magnitude of, or greater than, the gravitational force can keep the star buoyed up against gravity. If the force is short-range, at some critical mass gravity will overcome the repulsion and collapse will ensue; but the critical mass will be sensitive to the parameters of the repulsion. As the gravitational force is so weak (10^{-39} times the electromagnetic interaction between protons, for example), the assumption of even an extremely tiny force can have a dramatic effect. This is what Colpi *et al.* observe and develop in a full general-relativistic calculation.

Whatever the details of this calculation, the quantities and orders of magnitude involved are themselves illuminating. To set a baseline, the authors first assume non-interacting bosons. The radius, R, of a relativistic star (one near gravitational collapse, so that gravitationally bound particles have velocities near the speed of light) of such bosons will be of the order of the Compton wavelength of a single relativistic boson. Thus $R \sim 1/m$ in natural units, where m is the mass of the boson. (In natural units Planck's constant and the speed of light are taken to be unity. In such units, with masses in GeV, roughly the proton mass, lengths are approximately in femtometres.) But as the orbits are relativistic, R must be near the Schwarzchild radius GM, where G is Newton's gravitational constant and M is the mass of the 'star'. Combining these equations we find $M \sim 1/Gm$. Now we can replace G by its expression in terms of the Planck mass — the mass M_p such that the gravitational self-coupling GM_p^2 is roughly unity, and we find that the mass of the 'star' is $M \sim M_p^2/m$. This reproduces the bones of Ruffini and Bonazzola's result. The mass density of such a star would be M/R^3, or $M_p^2 m^2$, a gargantuan 10^{38} times nuclear density for 1-GeV bosons.

The first appropriate expression of Colpi *et al.* for the energy density of the interacting bosons, if ϕ is the boson field (or wavefunction, to use a more familiar name), is $m^2\phi^2 + \lambda\ \phi^4$. Here the first term approximates the mass and kinetic energy of a relativistic particle and the second is the assumed contact interaction term, which would arise, for example, if the bosons repelled one another by the exchange of a heavy intermediate particle. The question now is, what do we mean by 'small' interactions being important? By small we must mean that $\lambda\phi^4 << m^2\phi^2$, in other words that the interaction energy is much less than the kinetic or mass term. To extract λ from this inequality it is necessary to know the value of ϕ, the boson amplitude in the star. But the mass density of the star is $M_p^2 m^2$, as shown above, which must equal the boson-energy density. If λ is 'small' we can set the latter equal to $m^2\phi^2$. As a result, we see $\phi \sim M_p$, so that finally we get the limit on λ that $\lambda << (m^2/M_p^2) = Gm^2$. This means the self-interaction must be much weaker than gravity if its effect is to be negligible.

The next problem is to develop the theory with large λ, but the problem (S. L. Shapiro, personal communication) is that there are two dimensionless quantities to deal with in the calculations: λ and m/M_p. Hand-waving arguments thus fail to give any guidance, and it is necessary to solve the full relativistic equations. The resulting maximum boson star masses work out, for 'large' λ, to be around $M \sim 0.1 M_\odot \lambda^{1/2}/m^2$, with $M_\odot$ the mass of the Sun, rising as expected with the repulsion but falling as the square of the boson mass.

This is not the end of the story, however: refinements are still needed. Colpi *et al.* would like to deal with Higgs bosons, key elements of current unified field theories which in the empty (pre-inflationary) vacuum have an energy density much like that assumed above but with an attractive term (λ is negative). This leads to a condensation of Higgs field in the vacuum at the minimum of the energy density function (rather like the transition to superconduction in a superconductor). The condensate shortens the range of the weak interaction, giving the intermediate vector bosons the corresponding mass. Oscillations of the Higgs density are also allowed, and the corresponding waves are, in principle, observable, through wave–particle duality, as Higgs 'particles'. These are bosons, and could form boson stars — Higgs stars. Inside such stars, the Higgs field (which creates all particle masses) will vary, and the most intriguing physical possibilities arise. Unfortunately, Colpi *et al.* have so far been unable to solve the appropriate equations, which involve cubic as well as quartic terms, but work is in hand.

The group must also include the possibility of gauge forces such as colour and electromagnetism amongst the bosons. So far, Colpi *et al.* have assumed the particles to be neutral to a level equivalent to that of Ruffini and Bonizzola's neglect of possible λ terms. If the bosons were gauge charged, but on average neutral, the result would be an attractive force and the bose stars would undoubtedly collapse. So it will be necessary to consider non-neutral stars, those with a net gauge charge. Work on such stars, including the full Higgs terms, is now under way (S. L. Shapiro, personal communication). □

Robert Walgate is European Correspondent of Nature.

Plant molecular biology

Tailoring crop improvement

Chris Lamb and Leona Fitzmaurice

GENE manipulation techniques are likely to result in great improvements in crop production. Successful first steps towards introduction of disease resistance and herbicide resistance in plants were reported at a recent meeting*.

Cross-protection, inoculation of plants with mild strains of viruses, is often used to reduce losses due to more virulent strains, such as tobacco mosaic virus. To elucidate the molecular mechanism of resistance to tobacco mosaic virus infection, a chimaeric gene containing a cloned complementary DNA of the virus coat protein under the control of the cauliflower mosaic virus 35S transcript promoter has been introduced into tobacco cells on a tumour-inducing plasmid of *Agrobacterium tumifaciens*. The transfected plants produce high levels of the virus coat protein (R. Beachy, Washington University, St Louis; ref. 1). Protection from the virus, manifested as a marked delay in the appearance of symptoms, is dosage-dependent and can be partially overcome by inoculation with naked viral RNA. The number of primary sites of infection may be reduced when the virus is encapsulated by the coat protein, and the limited amount of protection to naked viral RNA suggests that there is a second mechanism affecting viral replication or spread. It will be of interest to see whether this strategy for engineering protection will be useful in other plants.

There has been progress in improving resistance of some plants to the broad-spectrum herbicide glyphosate, which inhibits the enzyme 5-enolpyruvyl shikimate 3-phosphate synthase (ESPS). A mutant bacterial enzyme which is insensitive to glyphosate inhibition confers some resistance to the herbicide when expressed in the cytoplasm of transgenic plantlets[2,3]. The plant enzyme is chloroplastic, and by expression of a chimaeric gene comprising transit peptide sequences of the small subunit (*ssu*) gene of ribulose bisphosphate carboxylase fused to the glyphosate-insensitive bacterial gene, the resistant ESPS can be made to accumulate rapidly in chloroplasts of transgenic plantlets (G. della Cioppa, Monsanto, St Louis; ref. 4). These plantlets are 1,000-fold less sensitive to the herbicide than those in which the transgene product accumulates in the cytoplasm.

Transgene regulation and protein targeting will be crucial for effective engineering, and new data indicate that 5′ flanking regions are not the only place where *cis*-acting gene regulatory sequences act. In petunia, the ribulose bisphosphate carboxylase small subunit is encoded by a family of eight genes that exhibit markedly different levels of expression. Analysis of the expression of transgenes comprising 5′ flanking, coding and 3′ flanking regions of different *ssu* genes indicates that sequences within the gene and 3′ flanking region have a major impact on expression (C. Dean and J. Bedbrook, Advanced Genetic Sciences, Inc.). Similarly, in chimaeric gene constructs containing the chloramphenicol acetyl transferase (*cat*) reporter gene and various promoters, the first intron from the maize alcohol dehydrogenase-1 (*Adh-1*) gene has a great impact on expression in electroporated maize protoplasts (V. Walbot, Stanford University). Expression of *cat* increased nearly 200-fold with the weakest promoter tested, 60-fold with the *Adh-1* promoter, and only 7-fold with the cauliflower mosaic virus 35S promoter, the most efficient promoter tested.

Future prospects depend on the development of an increased repertoire of gene transfer methods and systems, as well as the identification of plant genes that are agriculturally significant targets for manipulation. It is striking that only one plant receptor has so far been cloned — the photoreceptor phytochrome, involved in morphogenetic and tropic responses[5,6]. Two isoforms of phytochrome have been identified in etiolated plants. This polymorphism, as well as differences identified between phytochromes in etiolated and in green plants[7], could have considerable functional significance. The different polypeptides seem to arise from differential expression of members of a small gene family, although alternate translation of a single transcript, as well as post-translational modification of the proteins, may also be involved (P. Quail, University of Wisconsin, Madison). The au^{w} mutant of tomato exhibits many features of etiolated plants, and protein blot analysis indicates that the level of a phytochrome polypeptide is very low in the mutant. This system offers the potential for studying the relationships between structure and function of this receptor by examination of the effects of specific mutations engineered *in vitro* on phytochrome levels in transgenic plants following re-insertion of the modified genes, and may provide opportunities for precise engineering of changes in day-length requirements and response to shading of target crop plants.

Arabidopsis thaliana, the common wall cress, possesses several attractive features (ref. 8 and discussed in *News and Views*[9]). Its generation time of 4–5 weeks, the ease with which it may be propagated in large numbers, its lack of repetitive sequences and small genome size (about 70,000 kilobase pairs)[10] are advantageous for the rapid identification and cloning of mutated plant gene sequences. C. Somerville (Michigan State University) produced three *Arabidopsis* plants resistant to sulphonylurea by screening 300,000 mutagenized seedlings[11]. In collaboration with J. Smith and B. Mazur (DuPont), a yeast acetolactase synthetase gene was used to identify the gene encoding the *Arabidopsis* enzyme in a genomic library generated from one of these sulphonylurea-resistant plants. Somerville also identified *Arabidopsis* mutants that are defective in lipid biosynthesis[12]. Mutants for many of the biosynthetic steps, including most of the desaturase steps, can be identified by gas chromatographic analyses of leaf tissue derived from mutagenized plants. When these data are combined with two-dimensional gel analyses of proteins, it will be possible to identify proteins that are differentially expressed in the mutants and to clone the genes of this pathway.

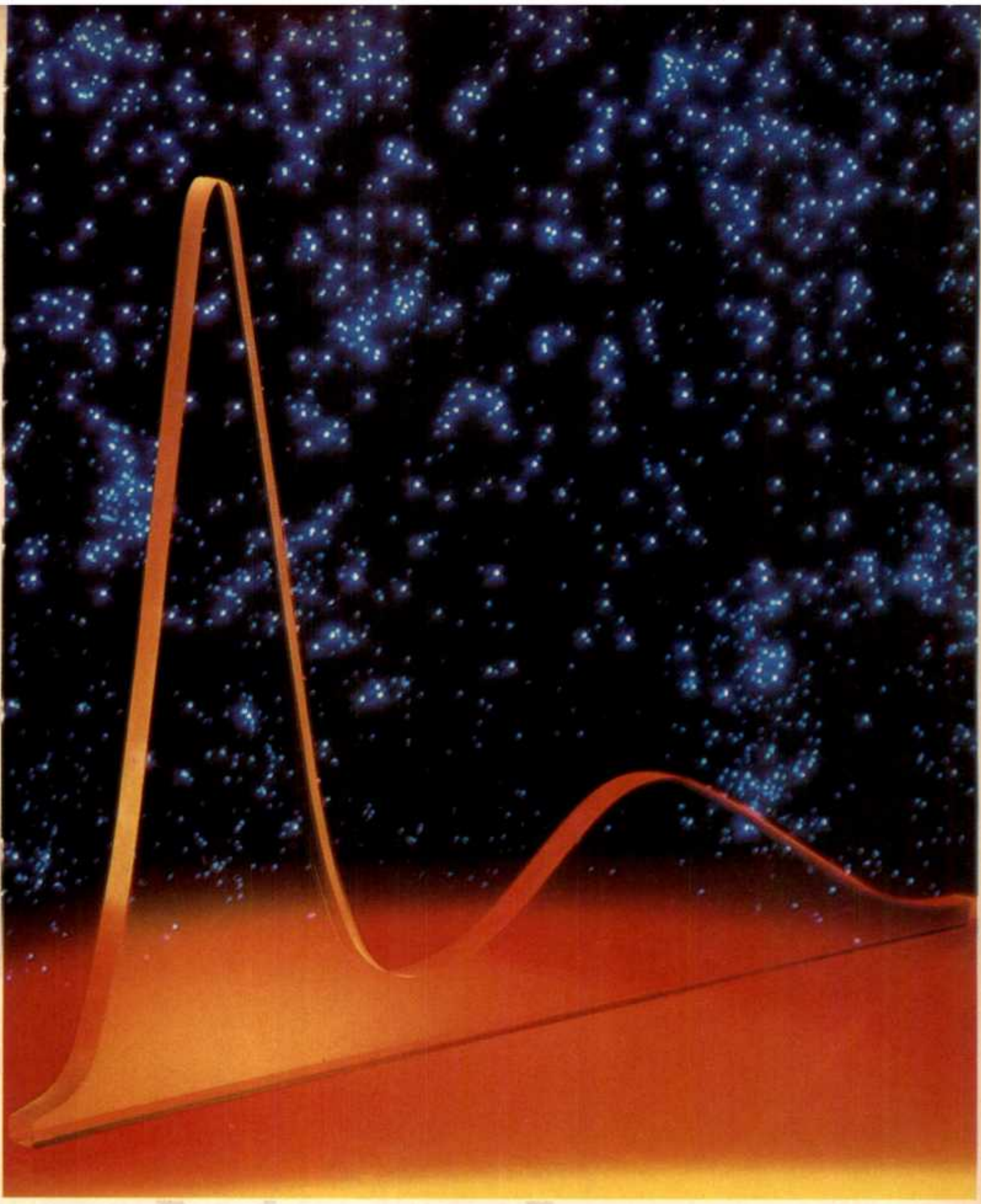
P. Ahlquist (University of Wisconsin, Madison) described the designed modification of brome mosaic virus for use as a vector[13]. Extremely high levels of expression can be achieved with this gene transfer system. Milligramme quantities of virion coat protein are produced in infections with the wild-type virus. With engineered constructs, *cat* activity already has been obtained which is five- to fifteen-fold greater than that achieved with tumour-inducing plasmid-based vectors, and may be further increased by control of messenger RNA and/or protein degradation. Future potential advantages of this viral vector are the broad range of monocot and dicot hosts as well as effective gene transfer to vegetative cells of perennial crop plants. □

1. Abel, P.P. *et al. Science* **232,** 738 (1986).
2. Comai, L. *et al. Nature* **317,** 741 (1985).
3. Shah, D.M. *Science* **233,** 478 (1986).
4. della Cioppa, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **83,** 6873 (1986).
5. Hershey, H.P., Colbert, J.T., Lissemore, J.L., Barker, R.F. & Quail, P.H. *Proc. natn. Acad. Sci. U.S.A.* **81,** 2332 (1984).
6. Hershey, H.P., Barker, R.F., Idler, K.B., Lissemore, J.L. & Quail, P., *Nucleic Acids Res.* **13,** 8543 (1985).
7. Tokuhisa, J.G., Daniels, S.M. & Quail, P.H. *Planta* **164,** 321 (1985).
8. Meyerowitz, E.M. & Pruitt, R.E. *Science* **229,** 1214 (1985).
9. North, G. *Nature* **315,** 366 (1985).
10. Leutwiler, L.S., Hough-Evans, B.R. & Meyerowitz, E.M. *Molec. gen. Genet.* **194,** 14 (1984).
11. Haugh, G.W. & Somerville, C. *Molec. gen. Genet.* **204,** 430 (1986).
12. Browse, J., McCourt, P. & Somerville, C. *Pl. Physiol.* **83,** 859 (1986).
13. Ahlquist, P. *et al. Proc. natn. Acad. Sci: U.S.A.* **81,** 7066 (1984).

Chris Lamb is in the Plant Biology Laboratory, the Salk Institute, PO Box 85800 and Leona Fitzmaurice is in the Plant Sciences Program, SIBIA Inc., PO Box 85200, San Diego, California 92138–9216, USA.

---

* *Tailoring Genes for Crop Improvement: An Agricultural Perspective* 24–28 August 1986, University of California, Davis, USA.



IN
THE NICK
OF
TIME.

With the Pharmacia NICK™-Column you don't need to compromise.

CHOOSE HIGH RECOVERY AND PURE SAMPLE. Your DNA recovery is at least 90% with less than 1% contamination of the nucleotides originally present in the sample.

CHOOSE SPEED AND CONVENIENCE. Your purified sample of nick-translated DNA is eluted in two minutes after column equilibration.

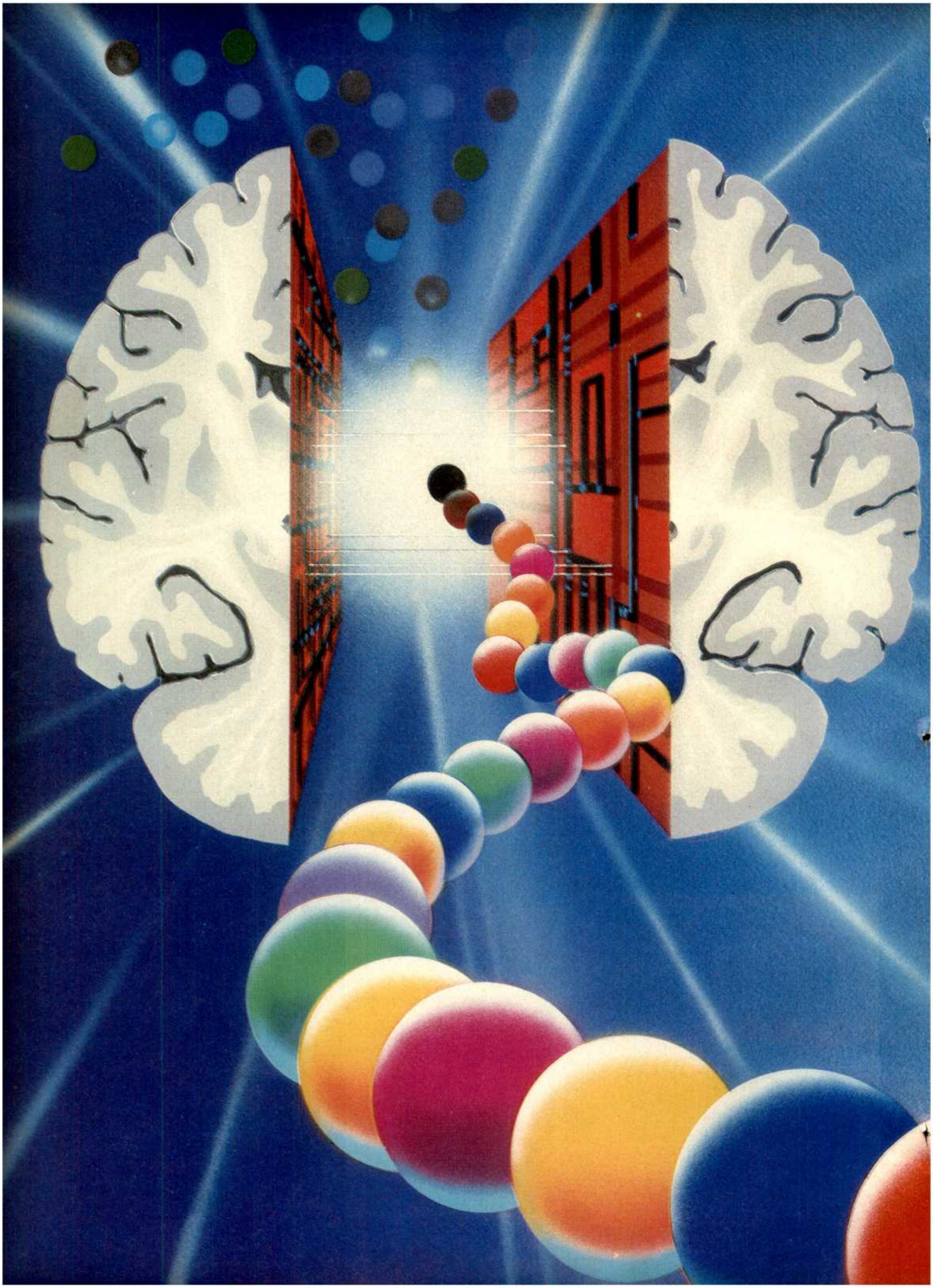
GET HIGH REPRODUCIBILITY. Pharmacia Sephadex® DNA Grade and our thorough quality assurance means that you get consistent results from batch to batch.

GET READY FOR IMMEDIATE USE. NICK-Columns are ready to use, prepacked with Sephadex G-50 DNA Grade, and a complete step-by-step protocol which comes with every pack.

Pharmacia

Molecular Biology Division
S-751 82 Uppsala, Sweden

Reader Service No.66



Put Du Pont's Knowledge in Molecular Biology to Work for You

Because we ourselves are biotechnology researchers, all Du Pont products are designed with an in-depth understanding of your needs. Each product is manufactured to meet high quality standards. This assurance of reliable performance can help save you precious research time and money... again and again.

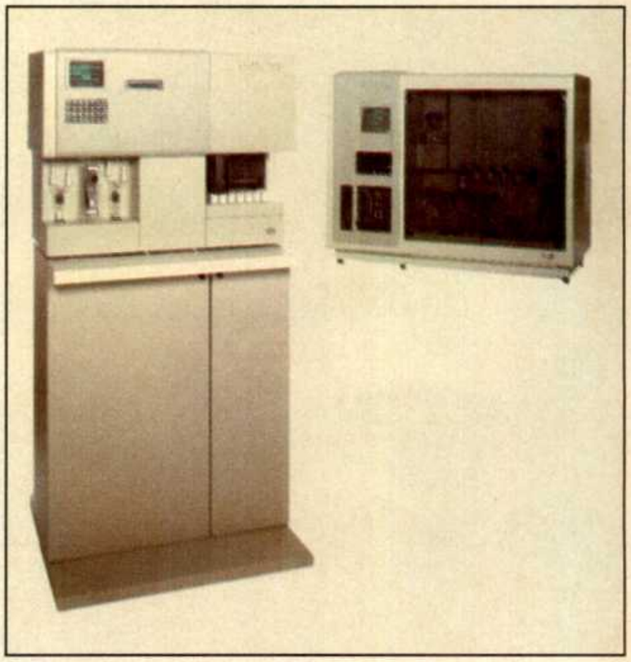
And, all are backed by a comprehensive support network of skilled technical experts and service representatives who are on call for you.

But you already know how well our knowledge works. We're the company providing over 1,000 molecular biology research tools, including widely used and highly accepted NEN® labeled chemicals such as ^{32}P nucleotides and ^{35}S amino acids; molecular biology research systems, such as the Nick Translation and *in vitro* Translation Kits; and NENSORB™ 20 nucleic acid purification cartridges. In addition, we provide SORVALL® Ultracentrifuges and ZORBAX® Bio Series HPLC Columns for molecular biology research.

Now, we bring you a new family of molecular genetic instruments to support you and your work in molecular biology.

Save research time in molecular synthesis with programmed strategy selection

COUPLER® 2100 Peptide Synthesizer—The new COUPLER® 2100 introduces Integrated Synthesis Logic (ISL). It considers both the amino acid being coupled and the existing peptide sequence, then selects and executes the most efficient coupling strategy... just as you would. The COUPLER® 2100 automates the formation of amino acid symmetrical anhydrides and active esters. And you can synthesize either 0.5 mmol or 2.0 mmol of peptide.



Du Pont's COUPLER® 2100 Peptide Synthesizer (left) and CODER® 300 DNA Synthesizer.

CODER® 300 DNA Synthesizer—The new CODER® 300 uses cyanoethyl phosphoramidite chemistry for state-of-the-art yields and purity, and has a flexible operating scale of 0.25, 1.0 or 5.0 μmol. Its unique solvent control system with proprietary low-volume/high-cycle reliable valves ensures superior mixing and highly efficient reagent use.

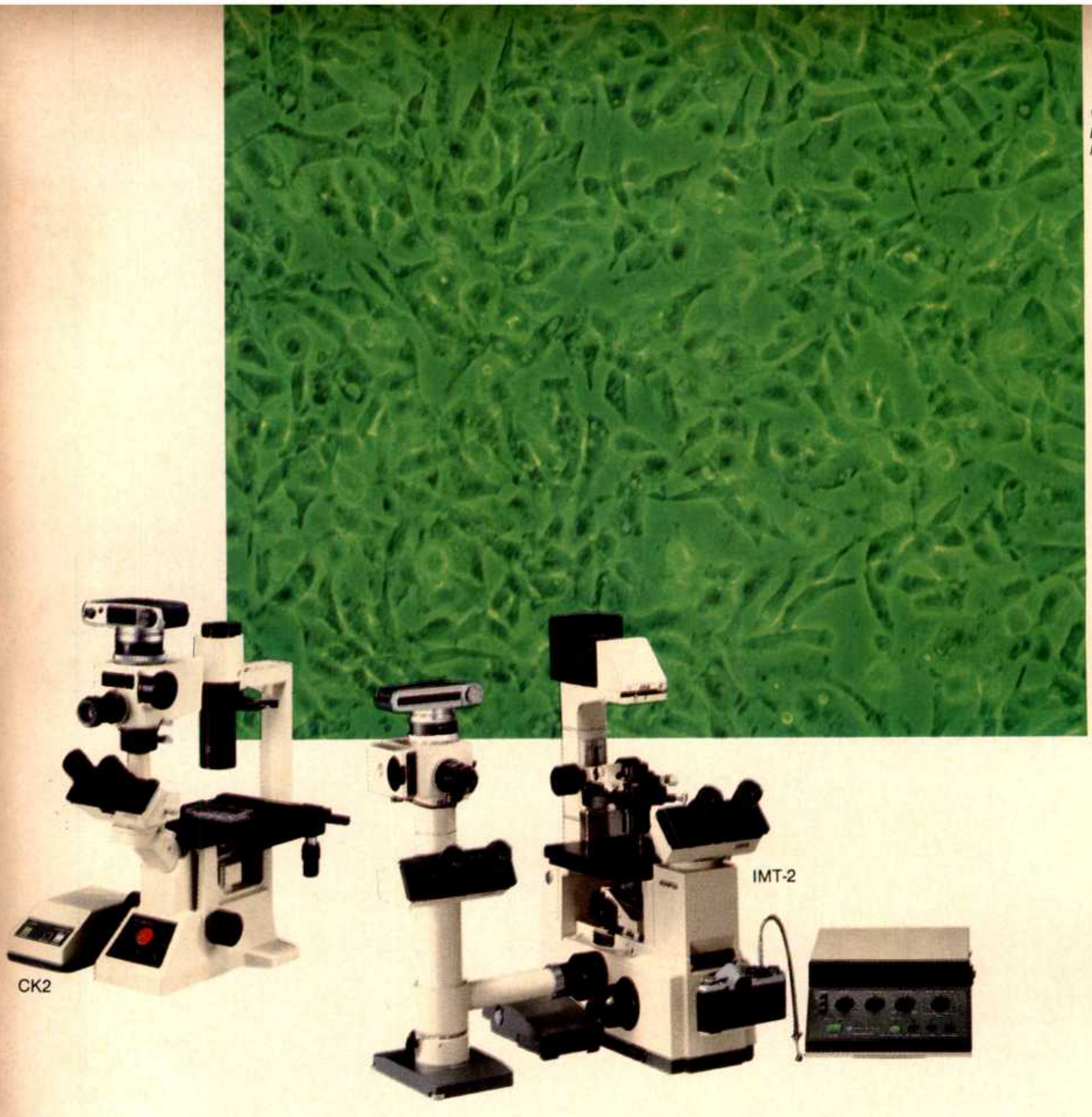
Call 1-800-551-2121 for more information, or write:
Du Pont Company, Biotechnology Systems,
BRML G-50554, Wilmington, DE 19898.

WORKING KNOWLEDGE
in biotechnology

DU PONT
REG U S PAT & TM OFF

Reader Service No.25

Visit DUPONT at ASCB Booth Nos. 598, 600 & 602



Vascular endothelial cell,
D Ach 10X PL,
NFK 3.3X LD

An unbeatable pair from Olympus

Olympus' new CK2 and IMT-2 inverted photomicroscopes redefine tissue culture microscopy.

Exceptional versatility and ease of operation are superbly blended into the Olympus CK2 and IMT-2 inverted tissue culture photomicroscopes.

The CK2 provides superior routine research functions in phase contrast microscopy and in photomicrography. And it's in a class by itself. Because it has a unique single-condenser system that is capable of handling a full range of magnifications from 4X to 40X—without changing condensers. Newly designed LB objectives capture crisper, sharper, aberration-free images for observation and photomicrography.

The sophisticated IMT-2 advanced research photomicroscope sets new standards of excellence. Its versatility is virtually unlimited, handling a full range of methodologies with minimal adjustments.

Superior high resolution optics, multiple viewing and multi-format photomicrography, a sextuple revolving nosepiece, and a wide variety of components make the IMT-2 ideally suited to tissue culture research and photomicrography.

Olympus — creating new standards in biomedical technology.

OLYMPUS®

OLYMPUS OPTICAL CO.,LTD.
San-Ei Building, 22-2, Nishi Shinjuku 1-chome, Shinjuku-ku, Tokyo, Japan

OLYMPUS OPTICAL CO. (EUROPA) GMBH
Postfach 104908, Wendenstrasse 14-16, 2 Hamburg 1, West Germany

OLYMPUS CORPORATION
4 Nevada Drive, Lake Success, N.Y. 11042-1179, U.S.A.

Reader Service No.165

CELL BIOLOGY TITLES FROM ELSEVIER

CELL STRUCTURE AND FUNCTION

Published by the Japan Society for Cell Biology (JSCB)
Distributed outside Japan by Elsevier Science Publishers

Editor-in-Chief:
Y. TASHIRO, *Department of Physiology, Kansai Medical University, Moriguchi-shi, Osaka 570, Japan*

Assistant Editor:
K. OMORI, *Kansai Medical University*

Aims and Scope
Cell Structure and Function publishes reviews and original research reports on the biology of cells from the molecular level to the level of cell differentiation and cell interaction. Preference is given to papers with an interdisciplinary approach using recent techniques of investigation in biochemistry, biophysics, molecular biology and structural organization. Papers with applied or clinical aspects will also be considered. A worldwide authorship is encouraged.

Subscription Information
1987 – Vol. 12 (1 vol. in 4 issues)
Price: US $ 106.00/Dfl. 265.00 including postage and handling.
US $ price is definitive.
Please apply to Amsterdam or New York address below.
ISSN 0386-7196

BIOLOGY OF THE CELL

Published under the auspices of the European Cell Biology Organization

Editor-in-Chief:
D. SANDOZ, *Centre de Biologie Cellulaire du CNRS, 67 rue Maurice Gunsbourg, 94200 Ivry sur Seine*

Editors:
R. CHARRET, *Paris/Orsay,* **M. DUBOS,** *Paris/Clamart* and **D. SZAFARZ,** *Paris/Orsay*

Aims and Scope
Biology of the Cell is an international journal publishing original reports, review articles and rapid communications concerning the structure and function of cells in the context of developmental biology, genetics, immunology and physiology. Papers about prokaryotes and viruses are also welcomed as well as those on new methodologies for cell study. Especially favoured are results obtained by complementary methodologies using recent techniques of investigation in cellular and molecular biology, which permit the comprehension of structural/functional relations in normal and pathologic cells. Papers devoted to *in vitro* and *in vivo* cell differentiation and cell interaction as well as those reporting original results on the organization of genetic material, cytoskeleton and membrane traffic are also of particular interest.
Special attention is given by the editor to the high quality reproduction of the illustrations.

Subscription Information
1987 – Vols. 59-61 (3 vols. in 9 issues)
Price: France: 1395 FF including postage, handling and V.A.T.
Rest of the world: US $ 219.00 including postage and handling
Please apply to Paris address below.
ISSN 0248-9000

CELL DIFFERENTIATION

The official journal of the International Society of Developmental Biologists (I.S.D.B.)

Editor-in-Chief:
L. SAXÉN, *Department of Pathology, University of Helsinki, Haartmaninkatu 3, SF-00290 Helsinki 29, Finland*

Associate Editor-in-Chief:
T. S. OKADA, *National Institute for Basic Biology, Myodaiji-cho 38, Okazaki, Aichi-ken 444, Japan*

Cell Differentiation is an international journal which serves its contributors by the rapid publication of novel research on **all** aspects of cellular differentiation in eukaryotes. Papers deal with molecular and structural aspects of cytodifferentiation and morphogenesis; gametogenesis and fertilisation; molecular and structural aspects of early development; cell and tissue interactions; pattern formation; genetic control of differentiation; growth factors and differentiation factors; and regeneration.

Subscription Information
1987 – Vols. 20, 21 (2 vols. in 8 issues)
Price: US $ 266.00 including postage and handling
Please apply to Limerick address below
ISSN 0045-6039

TECHNIQUES IN THE LIFE SCIENCES

Techniques in the Life Sciences will provide comprehensive coverage of techniques used in each of the major disciplines in the Life Sciences, for example, Biochemistry, Physiology, Cell Biology. A section, comprised of a number of individual volumes, will be devoted to each of these disciplines.

CELL BIOLOGY SECTION

CELL BIOLOGY – VOLUME C3 (1987)

Study of Cell Cultivation
Editor: **E. Kurstak,** *Montreal, Canada*
In preparation.

CELL BIOLOGY – VOLUME C2 (1986)

Techniques in In Vitro Invertebrate Hormones and Genes
Editor: **E. Kurstak,** *Montreal, Canada*
US $ 70.00

CELL BIOLOGY – VOLUME C1 (1985)

Techniques in Setting Up and Maintenance of Tissue and Cell Cultures
Editor: **E. Kurstak,** *Montreal, Canada*
US $ 77.00

For subscription information, please apply to Limerick address below.

ELSEVIER

FL37A849

ELSEVIER SCIENCE PUBLISHERS,
P.O. BOX 211, 1000 AE AMSTERDAM,
THE NETHERLANDS

EDITIONS SCIENTIFIQUES ELSEVIER
91, Rue de Rennes, 75006 Paris, France

Journal Information Center
ELSEVIER SCIENCE PUBLISHERS
52 Vanderbilt Ave.
New York, NY 10017, U.S.A.

ELSEVIER SCIENCE PUBLISHERS
IRELAND, LTD.
P.O. Box 85, Limerick, Ireland

Reader Service No.19

TITLES IN CELL BIOLOGY FROM

· ACADEMIC PRESS ·

Gene Activity in Early Development

Third Edition

Eric H. Davidson

September 1986, 490pp., $49.50/£41.50 (UK only), ISBN: 0.12.205161.0

This new third edition reflects the ten years of technological progress since the last edition. Providing a unique blend of classical and molecular knowledge, it discusses all major embryonic systems from both a comparative and mechanistic point of view. In deriving overall interpretations of developmental phenomena, it brings into play all the disparate forms of evidence, including genetic, molecular, and cytological.

Transport and Diffusion Across Cell Membranes

Wilfred D. Stein

1986, 685pp., $49.95/£42.50 (UK only), ISBN: 0.12.664661.9 Pbk.
$79.50/£67.00 (UK only), ISBN: 0.12.664660.0 Hbk.

Enhanced by numerous references and citations of important experimental findings in the field of biological transport, this volume thoroughly covers the molecular and kinetic aspects of the movement of molecules and ions across cell membranes. Step-by-step treatment of methods, kinetics, and design principles makes this volume very accessible to researchers who need an up-to-date, reliable resource on membrane transport.

Cell Culture and Somatic Cell Genetics of Plants

Volume 3 Plant Regeneration and Genetic Variability

Edited by **Indra K. Vasil**

1986, 688pp., $89.50/£74.50 (UK only), ISBN: 0.12.715003.X

This volume addresses the problems and questions surrounding the culture and modification of the many major crop plants, their regeneration, and the use of genetic variability to breed new, improved varieties.

Cell Separation

Volume 4 Methods and Selected Applications

Edited by Thomas C. Pretlow II and Theresa P. Pretlow

December 1986, 368pp., $59.50/£49.50 (UK only), ISBN: 0.12.564504.X

Methods for separation of specific types of cells such as cancer, basophil, and lymphoid are included in this volume of the treatise, which is designed for use by those not expert in cell separation who are searching for appropriate methods.

Current Topics in Cellular Regulation

Volume 28

Edited by **Bernard L. Horecker and Earl R. Stadtman**

December 1986, 337pp., $75.00/£62.50 (UK only), ISBN: 0.12.152828.6

prices subject to change without notice

Angiogenesis

O. Hudliká and K. R. Tyler

November 1986, c.230pp., $40.00/£25.00 (UK only), ISBN: 0.12.360130.4

This volume is a unique synthesis of information on the many aspects of blood vessel growth (angiogenesis) reflecting the increase in interest that has occurred in this field following the discovery of many growth factors and the development of appropriate tissue culture techniques. This material will be of relevance to a wide audience, including both basic researchers and clinicians, summarizing current knowledge and presenting hypotheses for future investigation.

Contents: Introduction. Techniques for Investigation of Capillary Growth. Growth of Vessels *In vitro* and *In vivo* — General Aspects. Growth of Vessels During Pre- and Postnatal Development. Comparative Aspects of Differentiation and Growth of Microcirculation in Different Species. Growth of Vessels in Normal Adult Tissues. Growth of Vessels Under Pathological Conditions. Factors Involved in the Growth of Vessels. Factors in the Inhibition of Vessel Growth. Conclusions.

Dimensions in Aging

The 1986 Sandoz Lectures in Gerontology

Edited by **M. Bergener, M. Ermini and H.B. Stahelin**

December 1986, c.350pp., $45.00/£25.00 (UK only), ISBN: 0.12.090162.5

This volume forms the proceedings of the 1986 Sandoz Lectures in Gerontology at which papers were presented by eminent scientists from various domains of aging research. As such it provides a valuable record of current research in this field and will interest all those involved in biological, medical and sociological aspects of gerontological research.

Under New Editorship

Genetic Engineering *Volume 5*

Edited by **Peter W.J. Rigby**

1986, 112pp., $19.00/£11.50 (UK only), ISBN: 0.12.270305.7 Pbk.

"The fifth volume in the series Genetic Engineering is now available. The emphasis under the editorship of Peter Rigby has moved away from reviews on the techniques of genetic engineering towards articles reviewing key areas of biology which, as a result of these techniques have been reborn."

Ann Dell, New Scientist, September 1986

Contents: Oncogenes, *A. Hall.* The Regulation of Gene Expression in *Eschericha coli* and its Bacteriophage, *C.F. Higgins.* Genes of the Immune System, *M. Steinmetz.*

Genetic Engineering Volumes 6—8

Articles in preparation include: The Genomes and Genetic Engineering of Plant Cytoplasms, *D.M. Lonsdale.* Vectors for Plant Engineering, *C.P. Lichtenstein.* The Structure and Function of Plant Genes, *J. Messing.* Molecular Biology of Parasites Causing Human Disease, *J.G. Scaife.* Molecular Biology of Trypanosomes, *C.E. Clayton.* Genes Controlling Development in *Drosophila, M.E. Akam.* Production or Proteins in Animal Cells, *M. Bendig.* Site Directed Mutagenesis, *G. Winter.*

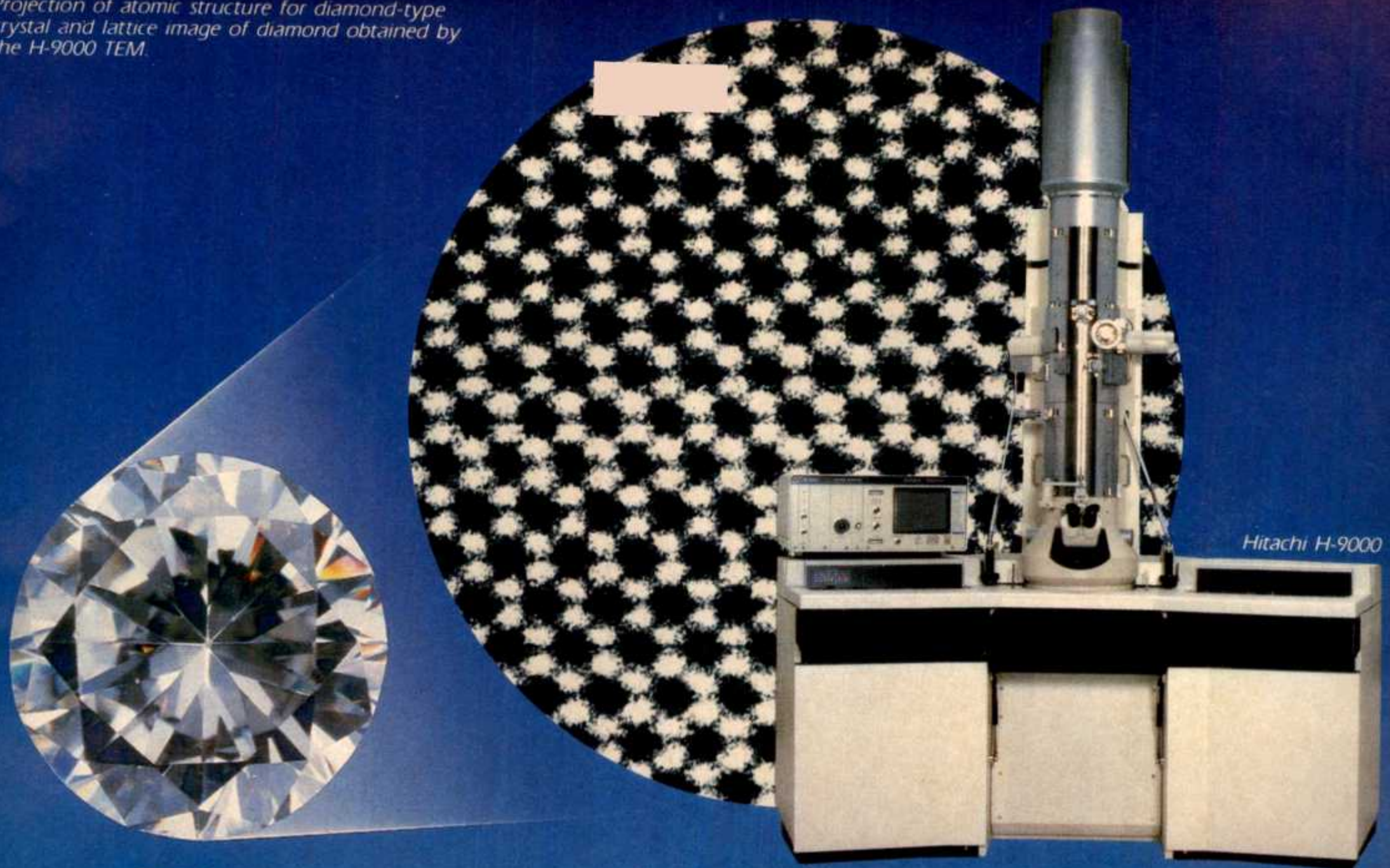
Interferon *Volume 7*

Edited by **I. Gresser**

November 1986, 224pp., $29.50/£16.00 (UK only), ISBN: 0.12.302256.8

24-28 OVAL ROAD, LONDON NW1 7DX, ENGLAND • ORLANDO, FLORIDA 32887, U.S.A.

Reader Service No.45



Projection of atomic structure for diamond-type crystal and lattice image of diamond obtained by the H-9000 TEM.

Hitachi H-9000

Clear as a Diamond

The Hitachi H-9000 300-kV TEM produces images with the stunning brilliance of a diamond.

The incredible penetration power and superb ultrahigh resolution of the Hitachi H-9000 300-kV transmission electron microscope has produced the world's first micrographs of the atomic structure of the diamond.

Dramatic Improvement Over Conventional 100-kV and 200-kV TEMs

Combining improved imaging resolution, a specimen penetration proportional to the electron velocity, an improved electron diffraction mode accuracy, and an improved spatial resolution for X-ray analysis, in which spread of beam penetration within a specimen is inversely proportional to the accelerating voltage, the Hitachi H-9000 TEM has produced stunningly clear micrographs of the diamond at the atomic level in remarkable detail.

This important achievement dramatically demonstrates the H-9000's ability to observe and analyze thicker specimens — a capability critical to the advancement of materials science, semiconductor technology, biomedical research, and other high-tech fields.

Compact Power

Based on Hitachi's extensive experience in million-volt microscopes, the H-9000 features a 10-stage acceleration system that provides ultrastable accelerating voltages of up to 300kV in an instrument of modest proportions. Enabling simple installation, use, and maintenance, the H-9000 easily fits in a normal-sized laboratory room, and operates just like a conventional 100-kV or 200-kV TEM.

Sophisticated Analytical Features

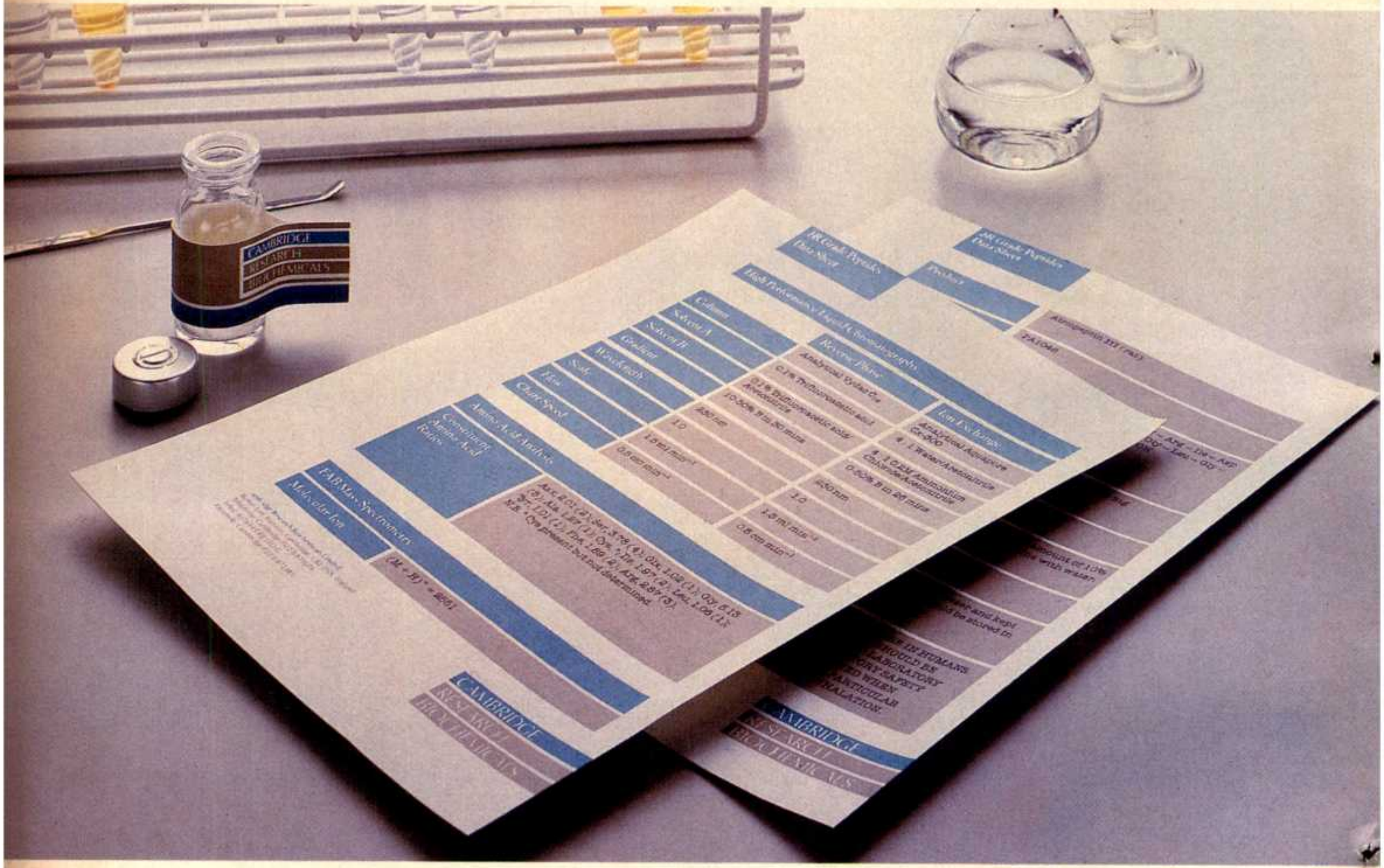
In addition to the conventional Be-window detector, a UTW detector may be mounted simultaneously, forming a dual EDX detection system. A special EELS using secondary aberration corrected three-dimensional focus spectrometry provides optimum analytical capability at 300kV. A TV monitor image intensifier system, SEM/STEM analytical applications, and a CBED fine probe add to the capabilities of the H-9000, one of the most powerful analytical microscopic systems available today.

Reader Service No.296

HITACHI

NISSEI SANGYO AMERICA, LTD. Tel: (415) 969-1100 **NISSEI SANGYO CANADA, INC.** Tel: (416) 675-5860
NISSEI SANGYO CO., LTD. England Tel: (0734) 66-4149 **NISSEI SANGYO GmbH** West Germany Tel: (0211) 450982—9
NISSEI SANGYO CO., LTD. Japan Tel: (03) 504-7111 **HITACHI, LTD., INSTRUMENT DIVISION** Japan Tel: (03) 214-3123

Free with every full analysis from Cambridge...



...a bR grade peptide.

The new range of 'biological reagent', (bR grade) peptides from Cambridge Research Biochemicals guarantees you a truly defined product for consistent, accurate results.

Setting the standard in peptide specification, each bR grade reagent is supplied by net peptide weight as determined by amino-acid analysis and enables standard solutions to be made directly from each vial.

Each product comes complete with a full analysis including amino-acid composition, high resolution HPLC and FAB mass spectrometry data confirming molecular weight, together with detailed storage and handling information.

Send now for your free copy of the new 'Cambridge bR Grade Peptides and Antibodies' catalogue which gives full details on over 100 products and custom synthesis services, and discover how bR grade peptides pass with honours at Cambridge.

CAMBRIDGE
RESEARCH
BIOCHEMICALS

UK. Cambridge Research Biochemicals Ltd., Button End, Harston, Cambridge, CB2 5NX, England. Tel: Cambridge (0223) 871674. Telex: 817694 CRB LTD G. Fax: 0223 872381
USA. Cambridge Research Biochemicals Inc., PO Box 58, 2005 Park Street, Atlantic Beach, New York 11509, USA. Tel: (516) 239 3831. Telex: 971398. Fax: 516 239 2782

BSF-2 is not just a differentiation factor

SIR—An interesting point has escaped Hirano *et al.*[1] in their report of the sequence of BSF-2, a B-cell differentiation factor. The point is that the sequence of BSF-2 is identical to that recently described by Haegeman *et al.*[2] of the State University of Ghent for a molecule called 26K. This molecule was originally described as a byproduct of interferon-β_1 (IFN-β_1) produced in human fibroblasts[3,4], and called IFN-β_2 after investigators at the Weizmann Institute detected antiviral activity associated with the molecule[4].

In both laboratories the primary structure of the 26K/IFN-β_2 molecule has been determined by molecular cloning of the complementary DNA[2,4,5]. But the question of whether the natural 26K/IFN-β_2 product possesses antiviral activity has remained a matter of controversy[3]. Hence, its true biological function has remained in question.

A first hint at an answer to this question came from the observation both in the laboratory of J. Content at the Pasteur Institute in Brussels and in our own laboratory that the mRNA for the 26K/IFN-β_2 molecule is induced in fibroblasts by interleukin-1 (IL-1)[6]. This suggested that the molecule might be one of the mediators of the multiple biological effects of IL-1. In an extensive search for biologically active substances released from fibroblasts by treatment with IL-1, Van Damme from our laboratory, together with J. Van Snick of the Ludwig Institute in Brussels discovered that fibroblasts exposed to treatments that induce appearance of the mRNA for 26K/IFN-β_2, release a potent growth factor for B-cell hybridomas and plasmacytomas[7,8]. In an effort to demonstrate whether this hybridoma/plasmacytoma growth factor (HPGF) is identical to the 26K/IFN-β_2 molecule, these investigators found that the mRNA for HPGF hybridizes to a cDNA probe of 26K/IFN-β_2[8]. Subsequently, HPGF was purified to homogeneity and, in collaboration with R. Simpson of the Ludwig Institute in Melbourne, its amino terminus was sequenced. This sequence matches exactly that predicted from the 26K cDNA clones and allows us to identify exactly the cleavage point between the signal peptide and the mature protein.

Recently there has been renewed interest in the possibility that the molecule may have direct antiviral activity[9]. Two cytokines, IL-1 and tumour necrosis factor (TNF-α), which are both able to induce the mRNA for the factor[6,9], also possess antiviral activity that can be neutralized by antisera against interferon-β. The question is whether these antiviral effects are mediated by IFN-β_1 or IFN-β_2. By using the B-cell promoting effect as an assay system, we have succeeded in preparing an antiserum that reacts with 26K/IFN-β_2 but not with IFN-β_1. This antiserum fails to neutralize the antiviral effect of IL-1 and TNF (J. Van Damme *et al.*, unpublished data). In contrast, sera that are specific for IFN-β_1 and do not neutralize the HPGF activity of 26K/IFN-β_2, do neutralize the antiviral effects of IL-1 and TNF[10]. This clearly demonstrates that IFN-β_2 does not play a role in the antiviral effect of IL-1 or TNF.

In conclusion, the substances known as BSF-2, IFN-β_2 26K and HPGF are one and the same. The substance certainly qualifies as a B-cell differentiation and growth factor but has little, if any, antiviral activity.

A. BILLIAU
The Rega Institute,
University of Leuven,
B-3000 Leuven, Belgium

1. Hirano, T. *et al. Nature* **324,** 73-76 (1986).
2. Haegeman, G. *et al. Eur. J. Biochem.* **159,** 625-632 (1986).
3. Content, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79,** 2768-2772 (1982).
4. Weissenbach, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77,** 7152-7156 (1980).
5. Revel, M. *et al.* in *The Biology of the Interferon System* (eds Stewart, W.E. & Schellekens, H.) 207-216 (Elsevier, Amsterdam, 1986).
6. Content, J. *et al. Eur. J. Biochem.* **152,** 253-257 (1985).
7. Van Damme, J. *et al. Eur. J. Immun.* (in the press).
8. Van Damme, J. *et al. J. exp. Med.* (submitted).
9. Kohase, M. *et al. Cell* **45,** 659-666 (1986).
10. Van Damme, J. *et al. Nature* **314,** 266-268 (1985).

Plant tumour induction

SIR—In Conrad Lichtenstein's otherwise excellent short comment on tumour induction by *Agrobacterium tumefaciens*[1], he omitted to mention the biological context of the bacterium. Wounding is a prerequisite for tumour formation, the basis for which is that wounded cells release inducers, such as acetosyringone, that activate the *vir*-region of the Ti plasmid of the bacterium, and thus facilitate its T-DNA transfer. *A. tumefaciens* occurs in the rhizosphere, and has been demonstrated to bind to plant cells. This implies that the bacterium is attracted to plants, but the biochemical basis for this has been left unexplained.

Intrigued, we set out to investigate chemotaxis in *Agrobacterium*. Using the methods of Adler[2,3], we have demonstrated that *A. tumefaciens* is capable of chemotaxis towards a range of phenolic plant wound exudates[4] including vanillyl alcohol. Crucially, we have discovered that chemotaxis towards the acetosyringone is determined by the Ti-plasmid. Interestingly, the chemotactic optimum occurs at 100 nM acetosyringone, some 100-fold lower than the concentration giving maximal *vir*-induction.

This suggests a biological system in which acetosyringone, at low concentrations, acts as a chemoattractant for *A. tumefaciens*. The bacterium thus moves up the concentration gradient towards the wounded plant cells. Then at higher concentrations, acetosyringone causes *vir*-induction, and ultimately T-DNA transfer.

CHARLES H. SHAW
ALISON M. ASHBY
MARTIN D. WATSON
Department of Botany,
University of Durham,
South Road, Durham, DH1 3LE, UK

1. Lichtenstein, C. *Nature* **322,** 682-683 (1986).
2. Adler, J. *Science* **166,** 1588-1597 (1969).
3. Adler, J. *J. gen. Microbiol.* **74,** 77-91 (1973).
4. Ashby, A.M., Watson, M.D. & Shaw, C.H. *FEMS Micro. Letts* (in the press).
5. Stachel, S.E., Messens, E., Van Montagu, M. & Zambryski, P. *Nature* **318,** 624-629 (1985).

An alternative view of the origin of life

SIR—In his *News and Views* item on the discovery of enzymatic activities in RNA molecules, Gilbert goes on to "contemplate an RNA world, containing only RNA molecules that serve to catalyse the synthesis of themselves"[1]. But what would be the source of energy in this RNA-run world? Another, and more basic question is, all these RNAs to what end? Are we interested in finding mere self-replicating machinery as an answer to the origin of life, or are we looking for a prototype of the self-propagating form as it exists in an organism?

At the simplest level, life is a massive molecular network that is able to gather a few substances to make usable energy and more molecules of different types that permit the replication of genetic material, which contains the information to run the network. The genetic material, the encoded proteins and the intermediary RNAs have received ready recognition from those engaged in the study of origin of life and its evolution; the tiny metabolites whose numerous linear, branched and cyclical conversions constitute the network, have received next to no attention. This is like embarking on the study of a kingdom and concentrating only on the king's family.

Enzymes can conceivably arise as the result of the formation of appropriate sequences of nucleic acids by random associations of nucleotides; but what function (and hence what selection value) can any of them have if there is no substrate to act upon? And, where does the substrate come from and what happens to the product? Meaningfulness demands a full pathway; but a pathway by itself is as meaningless as an enzyme in isolation. Pathways have to be able to interact with each other. The hypothesis of the formation of life by random associations of nucleotides and selection suffers from the need for prior and continual (for who knows when the gene encoding the proper enzyme might arise?) existence of the substrate molecules.

The 'central dogma' asserting the one-way transfer of information from DNA-to-RNA-to-protein has been challenged, apart from reverse transcriptase, by the proposition of reverse translation[2,3], and impressive evidence exists that anticodons and their corresponding amino acids show good sterochemical fits (see ref. 4). These findings deserve to be kept in view in any formulation on the origin of genes. It is conceivable that amino acids arranged around a prebiotically-formed stable molecule (all contained in the 'soup' trapped in the self-assembled lipid bilayer) became linked to form a miniprotein. If the miniprotein had catalytic activity for the stable molecule, this 'substrate' would be converted to a new molecule, setting in motion the formation of a pathway. Branching of pathways can be visualized as two distinct types of arrangement of amino acids around a molecule, generating two enzymes for the same substrate. In time the minienzymes could be reverse-translated and reversed-transcribed to minigenes. Their integration would form the genome, providing an explanation for the clustering of functionally related genes in bacteria.

The origin and evolution of life is the most basic problem of biology. Its elucidation demands the widest permissible latitude in thinking. Is it not time that we did away with the dictates of dogmas and labels like 'heresy'?

NARENDRA G. MEHTA
Biological Chemistry Division,
Cancer Research Institute,
Tata Memorial Centre,
Parel, Bombay – 400 012, India

1. Gilbert, W. *Nature* **319,** 618 (1986).
2. Mekler, L.B. *Nature* **215,** 481-484 (1967).
3. Cook, N.D. *J. theor. Biol.* **64,** 113-135 (1977).
4. Hendry, L.B., Bransome, E.D. Jr, Hutson, M.S. & Campbell, L.K. *Proc. natn. Acad. Sci. U.S.A.* **78,** 7440-7444 (1981).

The kingdoms of organisms

SIR—Having been concerned with the proper definition of eukaryote kingdoms[1-5], I wish to comment on the controversy over prokaryote kingdoms between Lake[6,7], Woese[8] and colleagues[9,10]. Three distinct questions are at issue: phylogeny, the chopping up of phylogenetic trees to make taxa, and the ranking of taxa. As Woese rightly insists, only phylogeny is a matter of objective science; the others, however, are not merely semantic[8] but questions of taxonomic judgement — that is, a mixture of art and science[11].

None of Lake's three trees[7] is inconsistent with the validity of Archaebacteria as a taxon. He confuses phylogenetics and classification in dismissing as irrelevant those characters he calls "three-and-one"; as if characters unique to all vertebrates are irrelevant to the validity of the taxon Vertebrata because they are irrelevant to overall animal phylogeny. Such characters are the most important of all in defining taxa. The distinctiveness of the wall and membrane chemistries of archaebacteria and eubacteria and their congruence with rRNA and tRNA sequences justifies the subdivision of prokaryotes into two major taxa of equal rank, the Archaebacteria and Eubacteria. This classification does not "fit all trees"[7]: it would be invalid if either taxon were shown to be polyphyletic, which Lake has not done.

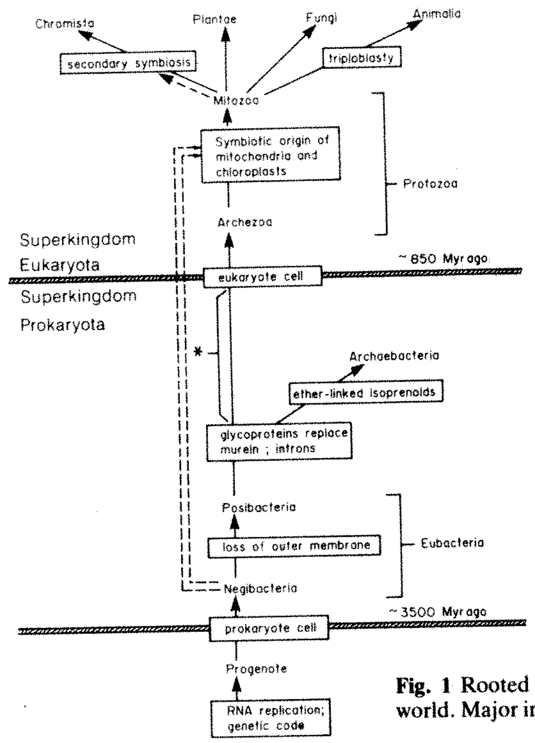


Fig. 1 Rooted phylogenetic tree of the living world. Major innovations shown in boxes.

The characters favoured by Lake are relevant to the subdivision of Archaebacteria into lower taxa and to which archaebacterial phenotype is closest to their nearest eubacterial relative. But his discussion of the distribution of the arginine deiminase pathway is flawed by ignoring evolutionary loss: repeated independent loss of characters in different lineages is common in evolution and is the most parsimonious explanation for the patchy distribution of arginine deiminase in bacteria. Because loss of enzymes is easy, one must not give too much taxonomic weight to the mere absence of a pathway of metabolite; many differences between Halobacteria and Methanobacteria can be attributed to the loss of ancestral proteins by the latter. The differences demand separation at least at the level of class but not at the level of kingdom or subkingdom. Similarities in RNA polymerases, antibiotic sensitivities and rRNAs make it sensible to rank Methanobacteria and Halobacteria as subphyla within a single phylum (=division), for which I propose the name Halomebacteria. I suggest that the sulphur-dependent archaebacteria ('eocytes') be ranked as a separate phylum (=division), the Sulfobacteria. This is compatible with all four 'prokaryote' trees.

The ranking of the taxa Eubacteria, Archaebacteria and Eukaryota has also been insufficiently carefully considered. Woese and Fox[13] referred to them as "urkingdoms"; this term, with no status or rank in the Linnean hierarchy, was chosen to avoid begging the question as to their taxonomic rank. Unfortunately numerous authors have carelessly substituted the Linnean term kingdom. But eukaryotes are universally subdivided into several kingdoms, commonly four[14], and more recently five[3-5]: Animalia, Fungi, Plantae, Protozoa and Chromista. Therefore Eukaryota must be ranked as a superkingdom[1,3], not a kingdom. The subdivision of the living world into two major structural cell types (eukaryote and prokaryote)[15] is in no way invalidated by recognition of the Archaebacteria and Eubacteria. This fundamental divergence, and their even more fundamental similarity in prokaryotic cell structure, can be most usefully represented by ranking Eubacteria and Archaebacteria as kingdoms within the superkingdom Prokaryota[12], rather than as subkingdoms[1,3].

The differences between halobacteria and photosynthetic eubacteria (all negibacteria) are so great that to place them in a 'photocyte' group[16] that excludes other archaebacteria is taxonomically meaningless. The affinities of the archaebacteria lie instead with Gram-positive eubacteria on the one hand and eukaryotes on the other[12]. The phylogeny shown in the accompanying figure is consistent with cell structure[12], with the 16S rRNA tree (if rooted where eubacterial phyla diverge)[8], and with the fossil evidence that eubacteria evolved 2,500 Myr before eukaryotes[17]. Lake's opinion[6,7] that eubacteria have a "fast clock" is counter to the fossil record which suggests they are the most conservative of the three groups, several times more so than eukaryotes.

Postulating a 'fast clock' or a 'slow clock' for a whole group does not explain the close bunching within, compared with the wide separation between, the three major lineages on the rRNA tree[8]. This pattern can be produced only by a temporary phase of accelerated rRNA sequence change during the divergence of the three lineages (on my tree at the time indicated by the asterisk, not seperately in three lineages diverging from the progenote as postulated by Woese).

Because it confuses phylogeny and classification, and is objectionable to eukaryote biologists, the misleading phrase 'three primary kingdoms' should be eschewed and replaced by the accurate 'three major lineages' when referring to the eubacterial, archaebacterial and eukaryotic branches of the tree of life. As there are two kingdoms of prokaryotes, and with the separation of the Chromista and Protozoa, five monophyletic eukaryotic kingdoms, we should speak of 7 not 3 kingdoms.

T. CAVALIER-SMITH

Department of Biophysics,
Cell and Molecular Biology,
King's College London (KQC),
25 – 29 Drury Lane,
London WC2B 5RL, UK

1. Cavalier-Smith, T. *BioSystems* **14**, 461 – 481 (1981).
2. Cavalier-Smith, T. *Biol. J. Linn. Soc.* **17**, 289 – 306 (1982).
3. Cavalier-Smith, T. in *Endocytobiology II* (eds Schwemmler, W. & Schenk, H.E.A.) 1027 – 1034 (de Gruyter, Berlin, 1983).
4. Cavalier-Smith, T. in *Progress in Phycological Research* **4** (eds Round, F.E. & Chapman, D.J.) 309-347 (Biopress, Bristol, 1986).
5. Cavalier-Smith, T. *Symp. Brit. Mycol. Soc.* **13** (in the press).
6. Lake, J.A. *Nature* **319**, 626 (1982).
7. Lake, J.A. *Nature* **321**, 657 – 658 (1986).
8. Woese, C.R. *et al. Nature* **320**, 401 – 402 (1986).
9. Zillig, W. *Nature* **320**, 220 (1986).
10. Lederer, H. *Nature* **320**, 220 (1986).
11. Simpson, G.G. *Principles of Animal Taxonomy* (Columbia University Press, New York, 1961).
12. Cavalier-Smith, T. *Ann. N.Y. Acad. Sci.* (in the press).
13. Woese, C.R. & Fox, G.E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5088 – 5099 (1977).
14. Whittaker, R.H. *Science* **163**, 150 – 160 (1969).
15. Stanier, R.Y. & van Niel, C.B. *Arch. Mikrobiol.* **42**, 17 – 25 (1982).
16. Lake, J.A. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 3716 – 3720 (1985).
17. Knoll, A.H. *Paleobiol.* **11**, 53 – 64 (1985).

How many reactor accidents?

SIR—It does not take advanced statistical calculations to conclude that there is a disturbingly high probability of another serious nuclear reactor accident in the next ten years, given that there have been two so far amongst 374 reactors with a total of 4,000 reactor-years experience. In their letter, Islam and Lindgren[1] are undoubtedly right to conclude from these data that the true probability is irreconcilable with the probability apparently claimed by reactor designers on the basis of 'technical risk assessment'. Their statistical arguments, however, are characterized by a failure to distinguish between probability and likelihood, which has led them to make unwarranted statements.

Assuming a Poisson process with parameter r (reactor-years^{-1}) they correctly find the probability of two accidents in T = 4,000 reactor-years to be $\tfrac{1}{2}(rT)^2e^{-rT}$. But they then treat this as an unnormalized probability distribution with parameter r, writing "If we want this to be a probability density, we have to normalize it". But one cannot turn any mathematical function into a probability distribution just by making it integrate to unity, and the function in question is a likelihood function for r, not a probability distribution, and cannot be integrated[2].

By treating the likelihood function as though it were a probability distribution, the authors have unwittingly adopted a Bayesian approach with implicit prior probability distribution dr for r. But even if one wished to admit the Bayesian approach here, no Bayesian would use this prior for a non-negative parameter since its integral obviously diverges and it distorts any analysis in favour of the larger values of r. The commonest Bayesian recommendation[3] is to adopt the arbitrary prior distribution $(1/r)$dr, which would have the effect of shifting the graph of the authors' Fig. 1 to the left, placing its maximum at r = 0.00025 rather than 0.0005.

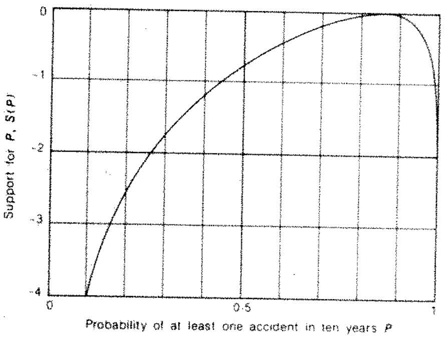


Fig. 1 The support for various values of the probability of at least one reactor accident in ten years, given two accidents in 4,000 reactor-years so far and the continued use of 374 reactors worldwide.

Islam and Lindgren then conclude from an integration of their probability distribution between two values of r that "This [the value 0.21] indicates that the probability that one significant accident will occur between 1,667 and 2,500 reactor years is 21%", a strange statement which presumably rests on an erroneous interpretation of their probability distribution for r (the caption to their Fig. 1 suggests this explanation).

The authors next compute the probability of at least one accident in t reactor-years, integrating out the unknown r in accordance with the Bayesian posterior probability distribution they have derived for it on the basis of the prior distribution dr and the two observations. They conclude "The probability that a reactor accident could happen during the next decade is 86%" (another strange statement since the probability that an accident could happen is manifestly 1; they mean will).

Aside, however, from these detailed criticisms, there remains the major issue of treating a likelihood function as though it were a probability distribution. Avoiding this issue, the analysis might proceed as follows.

As stated above, $\tfrac{1}{2}(rT)^2e^{-rT}$ is the likelihood function for the unknown parameter r in the light of two accidents in T = 4,000 reactor-years, and, as the authors say, the probability of at least one accident in t reactor-years is $P = 1-e^{-rt}$, and r is unknown. But rather than adopt a probability distribution for r we should use the above likelihood function for it and derive the corresponding likelihood for the unknown P by simple substitution: $r = -(1/t)\ln(1-P)$. So the likelihood for P is;

$$L(P) = \tfrac{1}{2}[-(T/t)\ln(1-P)(^2exp)(T/t)\ln(1-P)]$$

It is customary to work with the standardized log-likelihood, or *support*[2], which is;

$$S(P)l = 2\ln[-(T/t)\ln(1-P)] + (T/t)\ln(1-P) + 2 - 2\ln 2$$

(support is standardized by adding a constant so that the support function is zero at its maximum).

Figure 1 shows the support function for P with T = 4,000 and t = 10 × 374 reactor-years. The maximum-likelihood value for the probability of at least one explosion in ten years is seen to be 0.85 but, as is entirely to be expected with only two observations, there is great uncertainty about this value, the 2-unit support limits being 0.26 and 1.00. Islam and Lindgren's Bayesian analysis, by confounding likelihoods and probabilities, completely obscures this uncertainty, and consequently confers spurious scientific accuracy on what can only be an educated guess. All we can really say is that the probability of an accident in the next ten years is unlikely to be less than a quarter — a sufficiently alarming prospect not to detract from the main point they were making.

It may be tiresome to have to maintain these distinctions between likelihood and probability, but they are of especial importance when dealing with sparse data and infrequent events. Physicists usually use statistical theory in connection with large samples, where the finer points of statistical inference are relatively unimportant, but biologists often have to argue from quite small samples. It is no accident that the greatest impetus for the refinement of the logic of statistical inference has come from the biological sciences, and especially from human genetics, where the paucity of the data is matched only by the importance of the inferences. Nuclear accidents fall into the same category.

A.W.F. Edwards

Department of Community Medicine,
University of Cambridge,
Fenner's, Gresham Road,
Cambridge CB1 2ES, UK

1. Islam, S. & Lindgren, K. *Nature* **322**, 691-692 (1986).
2. Edwards, A.W.F. *Likelihood* (Cambridge University Press, 1984).
3. Jeffreys, H. *Theory of Probability* 3rd edn (Clarendon Press, Oxford, 1983).

Stunning whales

Sir—M.A. Taylor's comments on prey echolocation and sonic squid-stunning by whales (*Nature* **323**, 298, 1986) contain an incorrect discussion of the relation between the maximum detection distance by the whale (D) and the maximum distance at which potential prey can detect the whale's echolocating emissions (d). If for simplicity we neglect absorption, refraction, scattering by extraneous bodies, inhomogeneities, anisotropies, near-field and frequency-dependent effects, then the outward power-flux at distance r is $W/4\pi r^2$, and the returning power-flux experienced by the whale is $W\sigma/16\pi^2r^4$, where W is the emitted power and σ is the backscattering cross-section of the prey. If the minimum flux detectable by the whale is S and by the prey is s, we may find D and d from the equations $S = W\sigma/16.235^2D^4$ and $s = W/4\pi d^2$. The ratio d/D is $(WS/\sigma s^2)^{1/4}$. This expression can be made more illuminating by normalizing to prey-length, L, letting $\sigma = \alpha\pi L^2$ and $D = \beta L$, yielding $d/D = 2\beta(S/\alpha s)^{1/2}$. For any useful sonar system, the maximum prey detection range will be several prey-lengths ($\beta > 1$), and typically $\alpha < 1$.

Thus, for roughly equal hearing sensitivity, d/D will be greater than unity, and the prey will be able to take evasive action before detection by the whale. The ratio d/D is not 2, as indicated by Taylor's Fig. 1. This does not affect his fascinating discussion, but the relation given here permits additional variations of strategy in Taylor's "evolutionary arms race".

David C. Jolly

Box 931, Brookline,
Massachusetts 02146, USA

Body temperature and the specific heat of water

Sir—John Paul attempts to explain why the normal body temperature of homoiotherms is approximately 36 °C by its proximity to the temperature of minimum specific heat of water[1]. He states: "An organism functioning at this temperature will find it necessary to generate or dissipate the minimum amount of heat energy in order to maintain its temperature constant". But the rate of heat loss (and therefore rate of heat generation with which it is balanced) is equal to the temperature differential between body and environment times the heat transfer coefficient ("conductance"). This is independent of heat capacity, or heat content (specific heat times mass times temperature change). In fact, temperature would be maintained more easily if the specific heat and heat content were greater — that is, for a given amount of heat loss, the temperature decrease would be less.

Perhaps a better reason for maintaining body temperature considerably higher than the average ambient temperature is that this ensures that excess heat produced by the high metabolism of terrestrial mammals and birds can be dissipated by conductive, connective and radiative means which require no evaporative water loss[2].

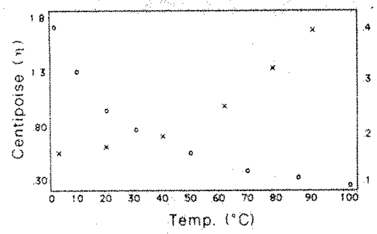
William A. Calder III

Department of Ecology
and Evolutionary Biology,
University of Arizona,
Tucson, Arizona 85721, USA

1. Paul, J. *Nature* **323**, 300 (1986).
2. Calder, W.A. & King, J.R. in *Avian Biology* Vol. 4 (eds Farner, D.S. & King, J.R.) 259–413 (Academic, New York, 1974).

Sir—We were intrigued by the proposal[1] that the nearly constant body temperature (around 36 °C) of many homoiothermic animals can be explained by the fact that the specific heat of water passes through a minimum around this temperature. The explanation offered was that "an organism functioning at this temperature will find it necessary to generate or dissipate the minimum amount of heat energy in order to maintain its temperature constant. From the point of view of the organism's energy economy this temperature is clearly the most efficient at which to function".

This explanation is appealing but it cannot be correct. The amount of heat required to maintain a given temperature is, of course, exactly equal to the amount of heat lost to the environment; this depends on many factors, such as the temperature difference, surface area and thermal conductivity of a body, but it is independent of its specific heat. Moreover, the smaller the specific heat, the larger will be the fluctuations in the temperature of the body, so that the problem of maintaining a constant body temperature becomes all the more difficult. Finally, the specific heat minimum at about 36 °C applies to pure water. Aqueous solutions generally shown no minimum in the specific heat, which usually decreases monotonically as the temperatue drops and is appreciably lower than that of pure water in the relevant temperature region. The specific heat of normal blood plasma is about 6 per cent lower, that of haematopoietic cells about 11 per cent lower[2].



Crosses, solubility; circles, viscosity.

What then is the explanation? We do not know, but most likely, a temperature of around 36 °C, if it is the product of natural selection at all, is selected because it corresponds to an optimal mix of properties. We consider here only two, viscosity and 'hydrophobic' effects.

The rates of many biological processes are limited by diffusion. As temperature rises, long-range structures in water are disrupted, the viscosity decreases, and the rates of diffusion-limited processes become faster. This factor by itself would appear to favour higher body temperatures. But disruption of the long-range structures in water will also lead to changes in hydrophobic effects, which are important in many vital processes, such as substrate binding, protein folding and bilayer membrane formation.

The accompanying figure shows the viscosity of water and, as a measure of hydrophobic forces, the solubility of benzene in water[3], both as a function of temperature. By inspection, a temperature of around 36 °C seems to be a reasonable compromise; high enough to give a low viscosity, low enough that hydrophobic molecules do not dissolve too easily.

Jack D. Dunitz
Steven A. Benner

Organic Chemistry Laboratory,
Swiss Federal Institute of Technology,
ETH-Zentrum, CH-8092 Zurich,
Switzerland

1. Paul, J. *Nature* **323**, 300 (1986).
2. Mendlowitz, M. *Science* **107**, 97 – 98 (1948).
3. Stephen, O.B.E. & Stephen, T. *Solubilities of Inorganic and Organic Compounds* (Pergamon, New York, 1963).

Reproductive failure in common seals

The table shown below was omitted from the letter by P.J.H. Reijnders on page 456.

Table 1 Number of participating, ovulating and pregnant seals in both experimental groups, during the season 1983–84

Group	1	2
No. of females	12	12
No. ovulating	12	12
No. pregnant	4	10

What is supersymmetry?

Michael Atiyah

Supersymmetry and its Applications: Superstrings, Anomalies and Supergravity.
Edited by G.W. Gibbons, S.W. Hawking and P.K. Townsend.
Cambridge University Press: 1986. Pp.481. £27.50, $49.50.

PHYSICISTS are an ambitious lot. Their aim is to understand the Universe and what makes it tick. What is matter? What are the forces that act on it? How did it start and where will it end? These, no less, are the monumental questions that theoretical physicists attempt to answer. So far, in historical terms, their track record in answering these questions is quite impressive. The two great triumphs of earlier centuries, Newton's theory of gravitation and Maxwell's theory of electromagnetism, remain outstanding examples of what a physical theory should be. Universal in application, and capable of simple and rigorous formulation in mathematical terms, they provide a firm basis from which to attack complex and realistic problems.

At the core of both theories are the notions of "particles" and "fields", that is of matter and force. Particles produce fields and fields act on particles in a perpetual duet. As physicists have delved further into the nature of this interaction, the theories that have emerged have become more and more complex, both conceptually and mathematically. Einstein altered our view of the separateness of space and time, and his theory of general relativity is much more sophisticated than Newton's, although equally rigorous in its mathematical formulation. Quantum mechanics involved an even greater revolution and its foundations are much less clear. "Elementary particles" have been emerging from the laboratories in ever-increasing numbers, and theories to explain them have been just as numerous.

In all this excitement and turmoil two main problems have always stood out. What are the ultimate "elementary particles"? And how does one reconcile general relativity with quantum mechanics? The first of these is closely tied to experimental discoveries and alters with each new generation of accelerators, while the second is a purely intellectual or mathematical problem because gravitation and quantum mechanics operate at quite different scales.

Quantum field theory aims to provide a theory of elementary particles consistent with quantum mechanics and based on a sophisticated unification of the notions of particle and field. The details of the theory may vary depending on how many kinds of particle one postulates. In fact, it is extremely difficult to construct any mathematically consistent quantum field theory. Many theories are plagued with "ultraviolet divergences" or "anomalies" and the mathematical study of these problems has become quite an industry.

Given all the difficulties, it is not hard to understand why, in the past couple of years, there has been such widespread enthusiasm amongst theoretical physicists for what are called "superstring" theories. These theories appear to offer the first realistic chance of a consistent account of nature which combines quantum field theory and general relativity. As the term suggests, these new theories are a combination of two different ideas, namely strings and supersymmetry, both of which have been investigated separately for some time.

In one form or another, symmetry has always played a prominent part in physics. Essentially, the requirement that a physical law be independent of the particular location or orientation in space or time means that the relevant theory has to admit the appropriate symmetries of space–time. Beyond this physicists have found that elementary particles tend to come in families, and the best explanation has again been couched in terms of symmetries, not of space–time but of some "internal" space.

Particles in quantum mechanics come in two different species, depending on their "spin". There are fermions, which have half-integral spin and obey the Pauli exclusion principle, and bosons which have integral spin and have symmetric wavefunctions. Physically they are fundamentally different but supersymmetry ignores this difference and puts them on an equal footing. Roughly speaking, a supersymmetric theory should allow the interchange of fermions and bosons. Despite its physical implausibility supersymmetry has many theoretical attractions, and the mere requirement that a theory be supersymmetric is so strong that it greatly limits the number of possible candidates. For these reasons it has been extensively studied.

String theory starts from the idea that the idealized picture of a point-particle should be replaced by that of a string. This also has been studied, off and on, for many years with varying degrees of enthusiasm and optimism. The subject really hit the headlines however when, in its supersymmetric form, it was discovered by John Schwarz of Caltech and Michael Green of Queen Mary College that a consistent anomaly-free theory could be constructed. This would in principle give a unified theory of elementary particles, incorporating gravity and quantum mechanics. Since then, there has been intense activity developing the several variants of the superstring theory and examining their consequences.

In 1985 a workshop in Cambridge, UK, was essentially devoted to these latest developments, and this volume records the various contributions. Many of the leading experts were present, and their talks give an authoritative view of the situation. While the material is highly technical, a few talks were addressed to the uninitiated, notably the introduction to strings by Edward Witten of Princeton.

Although superstring theories hold great promise and have generated renewed optimism about unified field theories, they are still very mysterious. The foundations are not at all clear, and surprises are probably in store for us. At the mathematical level, strings are difficult to understand, but it is already clear that they will involve many new techniques and ideas. It should be emphasized that the unification of forces involved in superstring theory takes place way beyond the realms of direct experimental verification. Thus the main tests of the theory are those of internal mathematical consistency. However, these are extremely severe. In fact physicists would like to believe that there is only one possible consistent solution. All they have to do is find it!

Perhaps the final word should be given to Witten, one of the high-priests of the subject, who, having explained the peculiar mathematics of superstrings, goes on to say:

> But the fact that these things work, that these seemingly bizarre rules give ways of computing things in quantum gravity, giving sensible results and finite answers and leading in many different directions to all kinds of beautiful areas in mathematics, is a very deep mystery — probably one of the deepest which has ever been encountered in physics. It is unlikely that a proper understanding of this mystery will be found either soon or simply. But it will be worth the wait. ☐

Sir Michael Atiyah is Royal Society Research Professor in the Mathematical Institute, University of Oxford, 24–29 St Giles, Oxford OX1 3LB, UK.

- Accounts of supersymmetry and associated topics are now reaching textbooks for advanced undergraduates and postgraduates. Four such books have appeared over the past few months and will be reviewed in a future issue of *Nature: Introduction to Supersymmetry* by P.G.O. Freund; *Introduction to Supersymmetry and Supergravity* by P. West; *Supersymmetry, Superfields and Supergravity: An Introduction* by P.P. Srivastava; and *Supersymmetry: A Decade of Development* edited by P. West.

Webs of intrigue

J.L. Cloudsley-Thompson

Spiders: Webs, Behavior, and Evolution. Edited by William A. Shear. *Stanford University Press: 1986. Pp.492. $55.*

WHEN, in his arbitrary way, Callimachus the Alexandrian announced that a big book is a big nuisance, he cannot have had in mind books like this one. It is, indeed, fairly large — nearly 500 pages in length — but far from being a nuisance and a bore it is an intellectual delight. Twenty-five years ago, worthwhile publications on the webs of spiders, their construction and functions, numbered a few dozen at the most. The bibliography of the present volume occupies over 38 pages, and the books and papers listed in it have mostly appeared in recent years.

The 13 chapters have been written by internationally recognized arachnologists, and it is no coincidence that they are either American or have done a lot of work in the United States — no longer is the study of spider behaviour confined to European naturalists and amateur zoologists. Four of the authors are ecologists who look at problems of web-building from the viewpoint of the interactions of spiders with the environment and with one another; eight are interested in behavioural biology (although half of them began as taxonomists); three others adopt a highly technical approach, focusing a laser beam on a single, vibrating silk thread; and one has found computers useful in monitoring the activities of a social species of spider. There are contributions on the selection of web sites, webs as sensory transducers, and the effects of the geometry of webs on their efficiency in catching prey: surprisingly, many orb-web spiders have reduced or even lost their web, so that they use up less rather than more silk.

By the secretion of chemical mimics of the sex-pheromones of Lepidoptera, certain spiders are able to attract male moths to a bolas of silk whirled by one of the legs. Other species construct narrow ladder-webs which attract large numbers of moths; as the prey slides down the elongated web, it sheds its scales until it has lost so many that its body is held firmly in place by the viscid threads. One nocturnal spider has acquired the anomalous behaviour of preying, not on flying insects, but on pedestrian arthropods. Its asterisk-shaped web has no sticky elements, and serves only to signal the presence of prey which is then wrapped with a swathe of silk so that the spider can administer a poisonous bite. Yet other species actively net prey in diminutive webs held between their legs.

Some spiders are kleptoparasites, living in the webs of larger species and stealing prey from them. More than a few attack their hosts as well. There is even a small parasitic spider that crawls over the bodies of larger species, feeding on the fluids that ooze from the prey of their hosts. Other of the more interesting topics developed in the various chapters are the monophyletic origin of the orb-web, and the evolution of web-building behaviour. It may surprise some readers to learn that jumping spiders (Salticidae), the quintessential visual predators of all land invertebrates, and equipped as they are with large and highly specialized eyes, have nevertheless evolved from web-building ancestors.

This is a book of puzzles, problems and paradoxes in spider behaviour, a synthesis of much recent work carried out, often in extremely uncomfortable tropical conditions, by a number of young and enthusiastic research workers. It is not difficult to understand the reasons for their motivation. □

J.L. Cloudsley-Thompson, Department of Zoology, University College London, Gower Street, London WC1E 6BT, UK, is Emeritus Professor of Zoology in the University of London.

Debatable physics

Michael Redhead

The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics. Edited by P.C.W. Davies and J.R. Brown. *Cambridge University Press: 1986. Pp.157. Hbk £17.50, $29.95; pbk £6.50, $9.95.*

HISTORICALLY, arguments about the interpretation of quantum mechanics divide into three periods. First, from 1927 to 1935, there was the famous debate between Bohr and Einstein over Bohr's complementarity interpretation. Bohr considered that quantum mechanics had led him to a profound metaphysical discovery about the physical world and our epistemological relationship to it; roughly, he argued that a complete understanding of microphysical reality requires us to recognize the existence of incompatible (complementary) descriptions, couched in the language of classical physics, but relativized to the specification of the whole experimental context in which that particular aspect of reality is being displayed. This is the so-called Copenhagen interpretation. Einstein argued that such relativization signalled a limitation in the completeness of the quantum-mechanical description, a fact which he sought finally to demonstrate in the famous paper he wrote with Podolsky and Rosen in 1935.

Bohr refused to accept the reasoning in that paper, and such was his charismatic authority that the scientific community in effect declared him the winner and for the next 30 years Copenhagen orthodoxy reigned supreme. However there were some dissenters, notably David Bohm, who in the early 1950s produced a "completed" interpretation of quantum mechanics along the lines envisaged by Einstein. However, Bohm's interpretation involved violently nonlocal features, whereas Einstein's 1935 paper had been predicated on the assumption of no-action-at-a-distance.

The third period of lively debate stems from the mid-1960s, when John Bell showed that *any* "completed" version of quantum mechanics must exhibit nonlocality as a result of violating the now famous Bell inequality. Two matters have been the centre of recent interest. First, are there any hidden assumptions in Bell's argument in addition to locality which could be impugned if the inequality is violated? And second, is the inequality actually violated experimentally? A succession of positive answers to the second question (after some initial indecisiveness) has culminated in the recent work of Aspect and his collaborators in Paris. With regard to the first question it is now largely accepted that "completed" versions of quantum mechanics do involve some feature of nonlocality, but of a rather subtle kind that does not for example allow the construction of "Bell" telephones for superluminal transmission of information!

The present book originated in a programme for BBC Radio 3 in which Paul Davies interviewed eight physicists with widely differing views on the interpretation of quantum mechanics. The most diehard defender of the orthodox position is Sir Rudolf Peierls. He announces: "there is only one way in which you can understand quantum mechanics". Such dogmatism (which is also factually incorrect) is what a number of other contributors are reacting against. In particular, Bohm and Hiley each explain how the traditional view of an external reality, possessing all its attributes independently of observation, can be accommodated in their quantum potential approach. They fully recognize the holistic or nonlocal features arising in their interpretation, but argue that this does not contradict the special theory of relativity, understood as applying to statistical effects.

Aspect lucidly describes the special features of his own experiments, while John Bell explains how he arrived at his inequality by "thinking intensely all around these questions". Bell usefully emphasizes that Aspect's experiment is not directly concerned with settling questions about indeterminism, of God's playing dice with the Universe, but bears directly on the distinct issue of non-locality. It is true, however, that one's understanding of this nonlocality and its apparent conflict with relativity may look very different from a deterministic and an indeterministic point of view.

David Deutsch explains and defends his own version of the Everett many-universes interpretation, that reality "splits" every time we take a measurement. His claim that the existence of the resulting "parallel" universes could be demonstrated experimentally is interesting but undoubtedly controversial. John Wheeler in his interview explains why he has now abandoned support for the Everett approach — "too much metaphysical baggage", as he puts it. He now leans to the view that the existence of conscious observers is crucial to quantum observation, a view definitely not to be ascribed to Bohr. Indeed it is fair to say that the whole language of observers and observables that clutters up books on quantum mechanics is an unfortunate legacy of a long-outmoded positivist attitude to the nature of science, but which tends regrettably to invite the casual reader to incorporate an ineliminable subjectivity into the understanding of quantum mechanics.

John Taylor attempts to defend what he calls the statistical interpretation, that quantum mechanics is a theory about ensembles of identically prepared systems, and has nothing to say about the single system. It is quite true that the nonlocality difficulties do not arise in terms of statistics. But Taylor's discussion of the Einstein–Podolsky–Rosen experiment is very confused, leading one to believe that mirror-image correlations do not exist in the *statistics* of measurement results on the same spin-component of the two particles. For example, Taylor says that "We can't say in any particular case what that spin of the far away particle is from the measurement nearby". This is potentially misleading, because it might be taken to imply that we cannot predict what the result of a subsequent far-away measurement would reveal, and that is just wrong. Presumably Taylor means that we cannot say what the spin of the far-away particle was *before* the measurement on the nearby particle. This is much closer to Bohr's position than what most people would regard as the statistical interpretation.

In all, this book is a useful introduction to a variety of ideas and approaches about understanding quantum mechanics. It shows that there is no longer a regimented uniformity of view even among "respectable" physicists. That is a healthy sign: dogmatic metaphysical strait-jackets are best consigned to the scientific lumber-room. □

Michael Redhead, currently Professor in the Department of History and Philosophy of Science, King's College London (KQC), Chelsea Campus, Manresa Road, London SW3 6LX, UK, is Professor-elect of History and Philosophy of Science at the University of Cambridge.

Captured in clay — 15,000-year-old models of bison from Le Tuc d' Audoubert cave in the French Pyrenees; the animal in the foreground is 60 cm long. The picture is reproduced from Dark Caves, Bright Visions: Life in Ice Age Europe, *by Randall White, an account in words and pictures of the accomplishments of the inhabitants of Europe during the Upper Palaeolithic (35,000–12,000 years* BP*). The book is based on an exhibition at the American Museum of Natural History, New York, and is published by W. W. Norton. Price is $35.*

Photograph by J. Vertut

Change in the field

Subir K. Banerjee

Environmental Magnetism. By Roy Thompson and Frank Oldfield. *Allen & Unwin:1986. Pp.227. £35, $50.*

THE TITLE of Thompson and Oldfield's book may lead one to think that it deals with the magnetic field and its fluctuations in the human environment. This, however, is not the case. The authors' respective backgrounds are in geophysics and geography, and over the past decade or so they have worked together on the magnetism of terrestrial materials — such as soil, and sediments deposited in lakes, peat bogs and oceans — with the purpose of reconstructing environmental and climatic changes. In spite of the complexities of natural magnetic minerals, the myriad ratios in which two or more such minerals can be present and the simplicity of the magnetic apparatus used, many of the stories told and the examples given constitute brilliant successes in scientific detective work.

Because of the many examples listed, the book will probably be of most use to non-magnetists interested in palaeoenvironmental change. It is well-structured, in that the first six chapters cover the basics of magnetism, magnetic minerals and the behaviour of the geomagnetic field, while the remainder deal with the applications to nature. The diagrams are many and of excellent quality, although a large number of them suffer from lack of acknowledgement of the source so that the reader will mistakenly think that they are either the results of the authors' own work or "given truths" in rock magnetism. For example, in Fig. 4.9 the authors purport to show the grain-size dependence of the ratio of susceptibility to saturation remanence for magnetite, but it is not made clear that these are the results of a thought experiment only. Similarly, on p.56 a non-magnetist not familiar with the literature may interpret the authors' hypothetical model of frequency dependence of susceptibility of very small magnetite grains to be proven, while a thorough test of such a model has yet to be made.

It is important to emphasize that nature's cooperation in sometimes providing a simple mixture of one strongly magnetic and one weakly magnetic mineral (for example, magnetite and haematite) has been most helpful to the authors in solving their magnetoforensic problems. In discussing one particular topic, the magnetism of lake sediments, Thompson and Oldfield themselves point out that "further progress, as distinct from the pragmatic application of the approaches illustrated above to new sites will depend on ... closer and more quantitative specification..." (p.123). I believe that the statement applies to most of the applications of magnetism to the environment, as described in this book. □

Subir K. Banerjee is a Professor in the Department of Geology and Geophysics, University of Minnesota, 310 Pillsbury Drive SE, Minneapolis, Minnesota 55455, USA.

FOCUS ON: CELLS

CELL AND MOLECULAR BIOLOGY OF THE CYTOSKELETON
edited by **Jerry W. Shay**
A thorough exploration of the cytoskeleton's molecular, biochemical, and structural properties with an emphasis on how they function in the living cell. Among the outstanding contributions in this volume are chapters on the structural aspects of intermediate filaments, molecular mechanisms controlling tubulin synthesis, probing the cytoskeleton by microinjection, and actin—a regulator of cell growth and differentiation.
0-306-42175-5/356 pp./ill./1986 $49.50 ($59.40 outside US & Canada)

VIDEO MICROSCOPY
by **Shinya Inoué**
A practical handbook for those getting started in video microscopy. The basics of both video and microscopy are reviewed and sufficient theory is provided to aid in the selection and use of video microscopy equipment. A list of hard-to-find pamphlets—along with the addresses of their sources—and a glossary—with relevant section numbers included for each entry—are among the many features that make this a valuable reference as well as a practical handbook.
0-306-42120-8/612 pp./ill./1986/$65.00 ($78.00 outside US & Canada)
Text adoption price on orders of six or more copies: $37.50

INTERACTION OF CELLS WITH NATURAL AND FOREIGN SURFACES
edited by **N. Crawford** and **D. E. M. Taylor**
One of the few interdisciplinary works to focus on the morphological and physiochemical aspects of cell-cell surface interactions and cell-foreign or natural surface interactions. Topics include electrokinetic properties of membranes, biotolerance, vascular grafts, sutures, and biocompatiability.
0-306-42373-1/proceedings/368 pp. + index ill./1986/$59.50 ($71.40 outside US & Canada)

Plenum
PUBLISHING CORPORATION

Plenum Publishing Corporation
233 Spring Street
New York, N.Y. 10013-1578
In the United Kingdom:
88/90 Middlesex Street
London E1 7EZ, England

Reader Service No.2

Other sides of the astronomer

Robert S. Westman

Nicholas Copernicus: Complete Works. Vol. III Minor Works. Edited by Paul Czartoryski. Translation from the Polish edition and commentary by Edward Rosen with the assistance of Erna Hilfstein. *Macmillan, London: 1986. Pp.380. £60.*

Few scientists and general readers are aware that Copernicus's first published work (1509) was a translation from Greek into Latin of a book entitled *Ethical, Rustic, and Love Letters* by an obscure ninth-century Byzantine writer named Theophylactus Simocatta. "Your natural beauty has faded and your good looks are approaching [the stage of] wrinkles", so opens letter No. 3. "But you try to give the impression of truth when you deceive your lovers with artificial cosmetics. Hearken to time, you hag, for in the autumn meadows are not the place for flowers." Etc. It is good fun and a fine way for a young humanist to sharpen his Greek. And, because Copernicus's Greek dictionary has survived, it is possible to see just how good his Greek was given the lexicographical tools at his disposal. Because every letter in his dictionary began with a capital letter, he could never be certain that he was considering a proper or a common noun. In letter 31, the peacock is described by Simocatta as "the Persian bird" who acquires "the arrogance of the Persians", but Copernicus produced instead "an active bird" that took on the "splendor of the energetic birds" (p. 8).

In 1517, Copernicus, a canon or administrative official of the Warmia chapter of the Church, composed a Latin treatise on the urgent matter of the debasement of coinage; two years later, at the behest of the West Prussian Estates, he prepared a German version that differs in certain respects from the first. "It is not the least advisable to introduce a new, good coinage", wrote Canon Copernicus, "while an old, debased coinage remains in circulation. How much worse was this mistake, while an old, better coinage remained in circulation, to introduce a new debased coinage, which not only spoiled the old coinage but, so to say, swept it away!" (p. 187).

The passages cited above are but two examples of the rich feast to be found among the "minor works" of Nicholas Copernicus, a figure otherwise known in the history of science for daring to propose that the Earth and the other "elements" beneath the Moon revolve annually around a motionless Sun. To be sure, astronomical writings are not missing from this volume: the *Commentariolus*, Copernicus's first formulation of the heliocentric hypothesis, probably written sometime between 1508 and 1514, and not published in Copernicus's lifetime; the *Letter Against Werner* (1524), which attacks Johann Werner's treatise on the motion of the sphere of the fixed stars but deferring to another occasion the author's own views on whether that sphere moves or not. In addition, one finds translations of all 17 extant letters written by Copernicus as well as a group of administrative documents. Copernicus had studied medicine at Padua; in one of his books, one finds a prescription for pills that promises to do the following:

> harmlessly eliminate whatever is excessive... postpone gray hair, which comes from corrupt humors...repress catarrh, stop a cough, relieve the congestion...alleviate gas in the stomach... sharpen the wits, strengthen and invigorate the nerves, preserve the teeth from decay...induce sleep...and purge gently [p.301].

Vast talents and international energies went into the preparation of this volume. Professor Paul Czartoryski, general editor of the *Complete Works,* deserves the highest praise for the difficult task of coordinating the Polish and American research teams. His job was not an easy one. Although the late Edward Rosen, assisted by Erna Hilfstein, produced the English translation, their work built significantly but with thin acknowledgement upon the Polish contributions. Few of the scholars mentioned in Czartoryski's detailed introduction, such as Jerzy Dobrzycki and Jerzy Drewnowski, are referred to in Rosen's notes; indeed, scholars will need to consult the Polish edition for material arbitrarily excluded by Rosen and Hilfstein.

On the other hand, a veritable firestorm of polemic fills Rosen's footnotes to the *Commentariolus* (pp. 91–126) where Noel Swerdlow (baptized as "Sw" in order to eliminate him from the Citation Index!) is roasted to a crisp for having had the temerity to translate anew in 1973 Rosen's supposedly canonical 1939 version of that same work. Rosen has now adopted some of Sw's language in this new translation — without acknowledgement. The principle of ignoring what you take from others while castigating them for their minor peccadilloes, real or imagined, seems here to have been raised to the status of a major editorial axiom. Copernican specialists will easily spot these off-the-text dog fights; general readers will probably miss them (without loss). Everyone will benefit from what is left over: a much richer picture of Copernicus as a Renaissance humanist, church astronomer and man of practical action. ☐

Robert S. Westman is a Professor in the Department of History, University of California, Los Angeles, California 90024, USA.

Announcing a new monthly journal that reports significant advances in basic cancer research and studies of growth regulation

ONCOGENE RESEARCH

AIMS AND SCOPE

Oncogene Research is a monthly journal dedicated to reports of significant research in the genetics, molecular and cellular biology of oncogenes, their products and the factors involved in the regulation of growth of normal and cancer cells **Oncogene Research** is intended to provide a vehicle for rapid publication of reports in this growing area of biology while maintaining a rigorous review process The journal will publish research papers, mini-reviews and short communications

EDITORS.

Claudio Basilico
Department of Pathology
New York University School of Medicine
550 First Avenue
New York, NY 10016 USA

Lionel Crawford
Molecular Virology Laboratory
Imperial Cancer Research Fund
Lincoln's Inn Fields
London WC2A 3PX UK

Carlo M Croce
The Wistar Institute
3601 Spruce Street
Philadelphia PA 19104 USA

François Cuzin
INSERM U273
Centre de Biochimie
Université de Nice
06034 Nice, France

Hidesaburo Hanafusa
The Rockefeller University
1230 York Avenue
New York, NY 10021 USA

Inder M Verma
The Salk Institute
P O Box 85800
San Diego, CA 92138 USA

Michael H Wigler
P O Box 100
Cold Spring Harbor Laboratory
Cold Spring Harbor, NY 11724 USA

CALL FOR PAPERS

We invite you to contribute to **Oncogene Research.** Papers can be submitted to the editor of your choice Decisions on submitted papers will be made as rapidly as possible, usually within four weeks from the date of receipt Publication of accepted manuscripts will be within three months Some of the various topics that will be considered for publication are

- Molecular structure of cellular and viral oncogenes and oncogene products
- Function of oncogene proteins
- Cellular transformation
- Expression of oncogenes during development
- Growth factors and their receptors
- Oncogenes, gene rearrangements and other mutations
- Recessive oncogenes and tumor suppressor genes
- Tumor viruses as mutagens
- Viruses associated with human tumors and cancers
- Oncogenesis in transgenic mice
- Evolution of oncogenes
- Regulation of cell cycle proliferation

First issue due March 1987
For further information closer to publication date, please contact

Harwood Academic Publishers
P O Box 786 Cooper Station New York NY 10276, USA • P O Box 197 London WC2E 9PX UK

Reader Service No 17

gauthier-villars

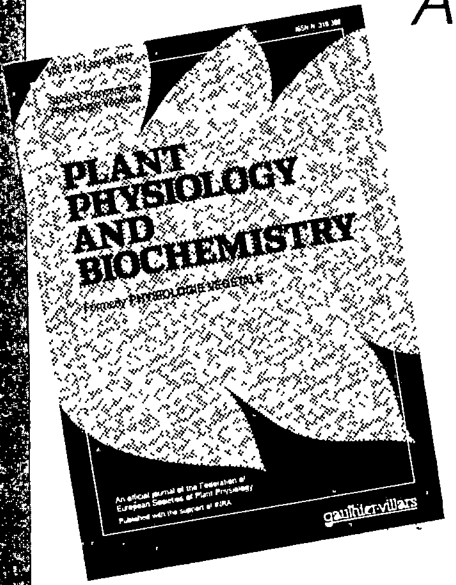
Starting in 1987 (vol. 25)

A NEW TITLE

A NEW COVER

A NEW FORMAT

for **PHYSIOLOGIE VEGETALE**



PLANT PHYSIOLOGY AND BIOCHEMISTRY

Editors
Jean-Claude KADER
Claude LANCE
Editorial office
Marie-France LAFORGE
12, rue Cuvier
75005 Paris, France

Manuscripts and all correspondence should be addressed to the Editorial office.
A guide to authors can be provided by the secretariat
- No page charges
- No membership needed for submission of Papers
- 50 reprints free

Annual subscription
1987 - vol 25 6 issues
978 FF (Export) - 730 FF (France)
Specimen copies sent on request

CDR - Centrale des revues
11 rue Gossin
92543 Montrouge cedex
France

Plant Physiology and Biochemistry publishes original theoretical, experimental and technical contributions in the various fields of Plant Physiology at diverse levels of integration (molecular, subcellular, cellular, organ, whole plant, environmental).

Plant Physiology and Biochemistry primarily publishes original experimental results It also publishes Short communications, Reviews, Letters to the Editor, Hypotheses, Technical articles and papers devoted to computer methods
Papers are published in English (72 % in 1986) and French.

Plant Physiology and Biochemistry is a journal of the Federation of European Societies of Plant Physiology (FESPP) As such, it cooperates with the two other journals of the FESPP, ***Physiologia Plantarum*** and ***Journal of Experimental Botany,*** through regular meetings of the Editors, exchange of expertise in reviewing the papers, common standards of presentation, etc

Plant Physiology and Biochemistry is published by the French Society for Plant Physiology.
The Society is open to all interested persons Application for membership and inquiries about Society activities should be addressed to the General Secretary
B. CAMARA, Société Française de Physiologie Végétale,
4, place Jussieu, 75230 Paris cedex 05, France.

Reader Service No 32

EUROPEAN JOURNAL OF medicinal chemistry

ELSEVIER IS PROUD TO ANNOUNCE THAT FROM 1987 ONWARDS, IT WILL PUBLISH THE **EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY** ON BEHALF OF THE **FRENCH SOCIETY OF THERAPEUTIC CHEMISTRY** AND OF THE **EUROPEAN FEDERATION OF MEDICINAL CHEMISTRY**

Created in 1965 as "Chimica therapeutica", this journal turned to its present title in 1974, when it had reached an European audience Now E J M C enjoys not only an European but an international reputation

AIMS AND SCOPE

The European Journal of Medicinal Chemistry publishes studies on all aspects of Medicinal Chemistry It provides a medium for publication of

- Original Papers
- Laboratory Notes
- Short Communications
- New Product News

The journal is intended for the use of medicinal chemists, pharmacologists, pharmacists, biochemists and physicians

- No page charges
- Instructions to authors are available at the editorial secretariat
- Manuscripts should be sent in triplicate to the Editor-in-Chief

Editor-in-Chief

Claude Combet Farnoux,
Faculte de Pharmacie
(Paris Sud),
5, rue Jean-Baptiste-Clement,
92296 Chatenay-Malabry,
France

Associate Editor

O Lafont, *Roven*

Editorial Advisory Board

P R Andrews, *Parkville*
B Belleau, *Montreal*
A Boucherle, *Grenoble*
J Cavalla, *Maidenhead*
J C Dearden, *Liverpool*
F G de las Herras, *Madrid*
J Elguero, *Madrid*
C R Ganellin, *Welwyn*
U Hacksell, *Uppsala*
H Horstmann, *Cologne*
R Jacquier, *Montpellier*
P Krosgaard-Larsen, *Copenhagen*
Ch Lespagnol, *Lille*
B Macchia, *Pisa*
Y C Martin, *Abbott Park, IL*
J Mathieu, *Romainville*
M Miocque, *Chatenay-Malabry*
R C Moreau, *Paris*
E Mutschler, *Frankfurt*
P Oehme, *Berlin*
H Pacheco, *Lyon*
J Poisson, *Chatenay-Malabry*
H Ramuz, *Basle*
M Robba, *Caen*
R Royer, *Paris*
J K Seydel, *Borstel*
H Tanida, *Osaka*
B Testa, *Lausanne*
J Thuillièr, *Paris*
H Timmerman, *Amsterdam*
J Tollenaere, *Beerse*
J G Topliss, *Ann Arbor, MI*
V Tortorella, *Bari*
P Tronche, *Clermont-Ferrand*
C G Wermuth, *Strasbourg*
W Werner, *Munster*
D T Witiak, *Columbus, OH*

Abstracted/Indexed in Chemical Abstracts, Current Contents (Life Sciences), Excerpta Medica (EMBASE), PASCAL-CNRS Databases, Science Citation Index

Subscription Information
1987 Volume 22 in 6 issues
Price (France) FF 810 00 including postage and TVA, (other countries) US $ 135 00 including postage
ISSN 0223-5234

ELSEVIER

Please send this coupon to
in the U S A /Canada
Journal Information Center
Elsevier Science Publishers
52, Vanderbilt Ave ,
New York, NY 10017, U S A

Rest of the World
Editions Scientifiques Elsevier
91 rue de Rennes
75006 Paris, France

Please send orders to the Paris address
Advertising rates are available on request

European Journal of Medicinal Chemistry

☐ Please enter me a 1987 subscription (payment enclosed)
☐ Please send me an invoice for a 1987 subscription
☐ Please send me a free sample copy (available February 1987)
☐ Please send information on the Progress in Medicinal Chemistry Book series
☐ Please add me to your mailing list

Name __________

Address __________

FL37A853

Reader Service No 21

Eight Issues in 1987

CELL MOTILITY AND THE CYTOSKELETON

In the past six years, **Cell Motility and the Cytoskeleton** has become a major forum for reporting all aspects of cell motility and its relationship to the cytoskeleton This publication provides its readers with an easily accessible source for the latest developments in the field In addition, the journal's 8¼″ × 11″ format accommodates the increased flow of papers and allows for high quality reproduction of micrographs

Highest priority is given to papers combining more than one of the following approaches dynamic morphology, ultrastructure, biochemistry, biophysics, and genetic and molecular biological analysis Theoretical and mathematical papers are encouraged if based upon, and consistent with, realistic data in the literature Those wishing to make contributions may contact the Editor Dr B R Brinkley • Department of Cell Biology and Anatomy • University of Alabama at Birmingham • Birmingham, Alabama 35294 Authors in Europe may choose to send their manuscripts to the European office Dr Jeremy S Hyams • Department of Botany and Microbiology • University College London • Gower Street • London WC1E-6BT, England

Topics of interest include

- molecular architecture, supramolecular structure, cell shape
- interactions of motile systems (e g , tubulin- and actin-based systems)
- interactions of cytoskeletal proteins, intranuclear movements
- non-motile roles of motile or cytoskeletal proteins
- biochemical, biophysical, and molecular aspects of cytoskeletal proteins and related control- or binding-proteins
- the pathology of motile behavior and the cytoskeleton
- nuclear-cytoplasmic transport and interchange
- action of drugs on motility and the cytoskeleton
- prokaryotic flagellar and gliding motility
- changes in organelle shape (e g , axostyles, costae), the nuclear matrix
- microtubule gliding and particle transport
- axonal transport, amoeboid motility, morphogenetic movements
- reticulopodial movement in foraminifera, radiolaria
- targeted intracellular particle transport, actin-based particle transport
- bulk cytoplasmic streaming (in protists, plant, animal, and fungal cells)
- movement of tissue cells in vitro or in vivo
- endo- and exocytosis, spreading of platelets, tissue cells
- the role of motile and cytoskeletal proteins in development
- mitotic movements (e g , kinetochore, centrosomes and particle)
- cytokinesis, muscular contraction, cytoplasmic contractility
- spasmoneme and myoneme contraction
- eukaryotic flagellar and ciliary movement
- the centriole, centrosome, and flagellar rootlet derivatives
- interaction of the cytoskeleton with the extracellular matrix

Volumes 7 & 8 1987, Eight Issues, $250 00, A personal rate of $80 00 is available to individuals Subscriptions at the personal rate must be paid by personal check, credit card, bank draft, or money order For all subscriptions outside the United States add $26 00 postage

For further information please contact the publisher:

ALAN R. LISS, INC.
41 EAST 11TH STREET • NEW YORK, NY 10003

Reader Service No 13

EDITOR-IN-CHIEF
B.R. Brinkley

FOUNDING EDITOR
Robert Day Allen

ASSOCIATE EDITORS
Charles J. Brokaw
Robert D. Goldman
Yukio Hiramoto
Jeremy S. Hyams
Peter Satir
James A. Spudich
Andrew G. Szent-Gyorgyi

EMERITUS
Noburo Kamiya

EDITORIAL BOARD
W. Brad Amos
Howard C. Berg
Kerry Bloom
Gary G. Borisy
Joseph Bryan
Monique Cachon
Don W. Cleveland
John S. Condeelis
Roger Cooke
Giulio Gabbiani
Benjamin Geiger
Ian R. Gibbons
Keith Gull
David Hartshorne
Wolfgang Haupt
Elizabeth D. Hay
Peter K. Hepler
Shinya Inoué
Raymond J. Lasek
Clive Lloyd
Eckhard Mandelkow
Mark S. Mooseker
N. Ronald Morris
Thomas D. Pollard
Lionel I. Rebhun
Joel L. Rosenbaum
Hikoichi Sakai
Joseph W. Sanger
Manfred Schliwa
S.J. Singer
Peter Steinert
Keiichi Takahashi
D. Lansing Taylor
Lewis Tilney
Richard Vallee
J.M. Vasiliev
Dieter G. Weiss
George B. Witman

US Supreme Court to review Louisiana appeal

Thomas H Jukes

An alliance of scientists and educational and religious leaders has converged to defend evolution from determined attack by the creationist lobby

THE appeal in the case Edwards *et al* appellants [Edwin W Edwards is Governor of Louisiana] versus Aguillard *et al*, appellees, will be heard by the US Supreme Court next week (10 December) The case pits creationism against evolution The action originated in a 1985 opinion of the US Court of Appeals, affirming the unconstitutionality of an act passed in Louisiana "Balanced Treatment for Creation-Science and Evolution-Science in Public School Instruction, Requiring the Teaching of Creation Science in Louisiana Public Schools Whenever Evolution is Taught"

The act requires that textbooks, library materials, and educational programmes include balanced coverage of both 'creation science' and evolution, that both evolution and 'creation science' be taught as theories rather than as proven scientific facts, that each local school board develop and provide to each public school teacher a curriculum guide on the presentation of 'creation science', and that the governor designate a panel of seven 'creation scientists' to assist local school boards in the development of these curricula The act argues that "creation science consists of scientific evidence and not religious concepts" and therefore should be taught in science classes The act states that its purpose was "For protecting academic freedom " However, the US Court of Appeals ruled that "the principle of academic freedom abjures state interference with curriculum or theory as antithetical to the search for truth The balanced treatment act is contrary to the very concept it avows "

Religiosity

The strategy used by creationists in writing the Louisiana act was developed as a response to their 1982 court defeat in Arkansas, where "the inescapable religiosity" of creationsism had been exposed by Judge Overton In the Louisiana act they kept mum about the religious nature of creationism and pretended that a viable 'creation science' existed

The Governor, in his officical capacity, together with the Attorney General and State Superintendent of Education of Louisiana, are appealing against the decision of the US Court of Appeals of the State The legislature passed the act by a large majority, and it appears to have a strong popular appeal in Louisiana The appellants are heartened because the Appeals Court ruled on a majority of only one seven judges dissented

The US Supreme Court procedure is that the appeal will be argued by lawyers for the appellant and appellees Permission has been granted for other interested parties to file *amicus curiae* briefs These will be added to the record to be read by the Court, but the lawyers who filed the briefs will not be allowed to present oral arguments on 10 December

FASEB's brief

The Supreme Court has received 16 *amicus curiae* briefs, most of them in support of the appellees As far as I know, the scientific and educational communities have never previously joined together on a legal issue to a comparable extent One brief for the appellees is joined by the Federation of American Societies for Experimental Biology with 28,000 elected members, nearly all with doctoral degrees Another brief was filed by 72 Nobel laureates and 17 state academies of science Probably this arousal of activism has come about because creationism threatens other natural sciences, such as astronomy, astrophysics and geology, in addition to biology A third brief for the appellees was filed by four clergymen

A typical statement by creationists is that they are "fully committed to the inerrancy of Scripture", as well as to "the strict creationist young Earth position taught in Scipture " This places the date of creation as about 10,000 years ago, which accounts for opposition by creationists towards any of the natural sciences that state that the Earth and the Universe are thousands of years old Creationists reject radiometric dating, by claiming that isotopic changes do not occur at uniform rates Similarly contrived arguments are used by creationists against any and all scientific findings that are in contradiction with a literal intepretation of the King James version of the Holy Bible

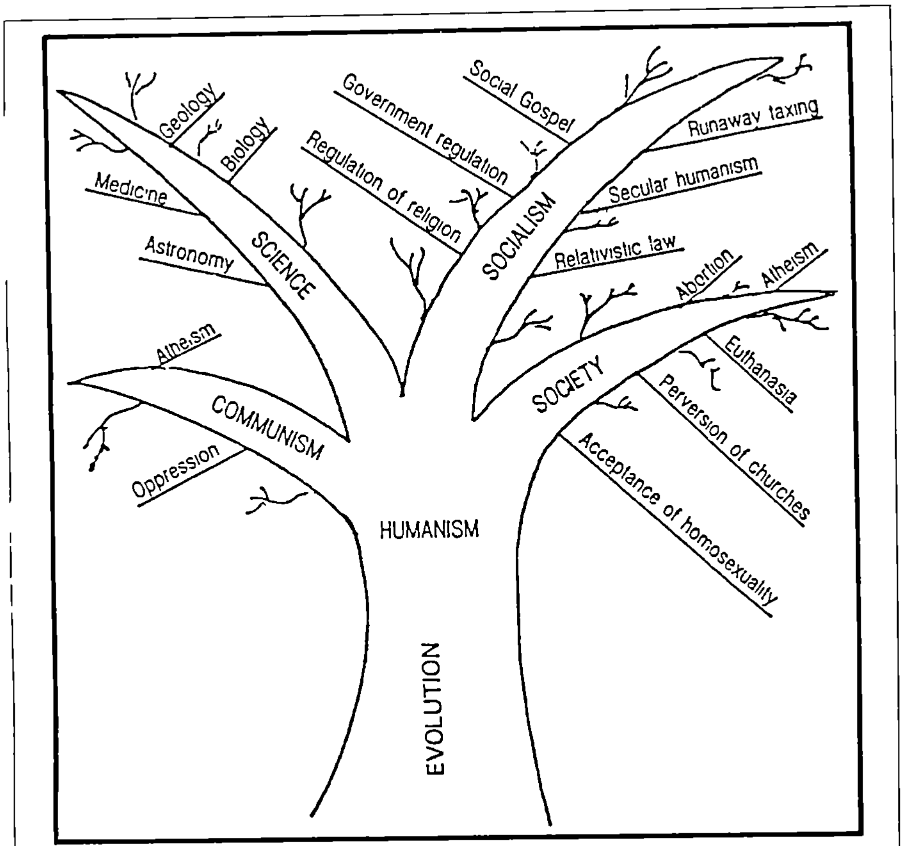
An irony is that the present case depends on legalities surrounding the establishment clause of the United States Constitution rather than on the validity of evolution Only one brief, that of the US National Academy of Sciences, says anything descriptive about evolution, in a discussion of divergence of cytochromes *c* in terms of the molecular evolutionary clock News of the current studies of proteins and DNA that reveal relationships of all living organisms with each other have not yet reached most people, and, of course, the knowledge has been deliberately excluded from most public school science textbooks by the successful efforts of creationists This exclusion is a victory for obscurantism It is therefore strange to read in *Nature* articles and letters affirming that creationism is a rather interesting sociological phenomenon, and one that should be treated sympathetically[1 2]

Four of the briefs are on behalf of the appellants, that is, for creationism, by the Christian Legal Society, the Rabbinical Alliance of America (joined by Congressman William E Dannemeyer, who is well known for his opposition to the pasteurization of milk), the Catholic League for Religious and Civil Rights (anti-abortion) and, as mentioned below, Concerned Women for America

The other 12 briefs are for the appellees, and usually emphasize the constitutional separation of church and state which the Louisiana Act would violate, also the attack on academic freedom presented by the act, and the threat to the teaching of other secular subjects, in addition to evolution, that is posed

"Anti-Christian"

In view of thepending Supreme Court case, a ruling by US District Court Judge Thomas G Hull in Greenville, Tennessee on 25 October, 1986, takes on added significance He ordered public schools in Hawkins County, Tennessee, to excuse fundamentalist children from attending classes where they would be exposed to books that their parents say promote themes which they regard as "anti-Christian " The parents objected to stories in the Holt-Rinehart-Winston books



“Evolution is the taproot which is feeding the oppressive, murderous and infidel directions we see gaining acceptance around us ” says Paul Bartz of Bible Science Newsletter

about “Cinderella” and “Macbeth,” because of their mention of magic, to *The Wizard of Oz,* because it portrayed a witch as good, and *The Diary of Anne Frank* because it suggested all religions are equal Ominously, another objection was to stories about dinosaurs if the creatures were said to be older than the Biblical account of the beginning of the world, presumably 10,000 years ago Judge Hull's decision included a statement that “the plaintiffs believe that after reading the entire Holt series, a child might adopt the views of a feminist, a humanist, a pacifist, an anti-Christian, a vegetarian, or an advocate of “one world government” He found that the plaintiff's objections were sincerely held religious convictions, thus apparently conceding that objections to almost anything should be upheld if under the guise of religious convictions The leading complainant also objected to reading about catholicism, which she said “could produce changes in my children's way of thinking—they could become confused”

It is of added significance that Michael Farris, the plaintiffs' chief attorney, is the author of the *amicus curiae* brief of the Concerned Women for America in the Supreme Court case In it, he argues that “the Louisiana law has the clear secular purpose of promoting academic freedom and of insuring that students learn the scientific evidence of both evolution and creation” This fallacious argument is based on the erroneous presumption that there is scientific evidence for creation, which evidence actually rests on literal acceptance of the Book of Genesis It is noteworthy that the Concerned Women for America, in one court action, object to *their* children reading about Cinderella, and in a second court action would foist creationist travesties of science on *all* children

Judge Hull also stated “This opinion shall not be interpreted to require the school system to make this option available to any other person or to these plaintiffs for any other subject ” Does this statement represent a disturbing challenge to the doctrine of *Stare decisis e non modere quieta*, which could thus be used against previous decisions that permit the teaching of evolution?

Another component that must be weighed in a Supreme Court case is the presence of William Rehnquist in his newly appointed role as Chief Justice Given the impact of the recent US Court decision in Greenville, what can we expect from him? It should be noted that the US District Court first threw out the case in December 1983, following which the parents appealed to the United States Court of Appeals in Cincinnati This court sent the case back to the lower court, holding that the parents' complaints deserved the hearing of a trial, with the result that the parents won Will the Supreme Court react to this precedent?

Taxation

The fantastic accusations against evolution by the creationists are seemingly endless The figure shown above accompanied an article written in 1983 by Paul A Bartz, the executive editor of *Bible Science Newsletter* The article included the sentence “Evolution is the taproot which is feeding the oppressive, murderous and infidel directions we see gaining acceptance around us ” Mr Bartz builds a powerful case for his opposition to evolution by alleging that it leads to “runaway taxing ” In a sense, this is quite true, in that runaway taxing did not exist before the appearance of vertebrates

It is more likely, however, that by “evolution” Bartz refers to the teaching of evolution rather than to its existence A sample of creationist “teaching material” for young children[4] contains a cartoon for each letter of the alphabet The letter *A* shows Adam and Eve surrounded by dinosaurs, *B* for Bible and brachiosaurus has Adam at his first job — naming the brachiosaurus *F* says that fossils were formed by the sudden inundation of the Biblical Flood, but “normally, when an animal dies, no fossil is formed” We are assured that “God put dinosaurs on the Ark” A chilling cartoon on *U* (for “unchanged”) implies that, just as the mockers of Noah were drowned, those who today jeer at church-goers will be destroyed by the vengeful Deity

Evolution started in Darwin, and grew with genetics The mechanism of inheritance had to provide for changes occurring during long periods of time Information in the genes is used every time a living organism develops from a fertilized egg cell, this event, because it happens so quickly, is far more remarkable than evolution The pervasive influence of biochemistry grew steadily, biochemists showed the universal nature of many metabolic pathways and showed that all living organisms had coenzymes containing the same remarkable molecules Specific procedures have firmly established evolution at the centre of the biological sciences, aided by population genetics, plate tectonics and computer sciences “Gaps in the fossil record” are irrelevant in the face of molecular evidence All lines of scientific inquiry show that the great apes are our closest relatives We are sorry that the creationists do not approve of their cousins, but, regardless of the Establishment Clause, this is the way it happened ☐

Thomas H Jukes is in the Department of Biophysics and Medical Physics, University of California, Berkeley, California 94720, USA

---

1 Marsden, G M *Nature* 305, 571-574 (1981)
2 Cavanaugh, M A *Nature* 315 185-189 (1985)
3 Bartz, P A *Bible Sci News l* 22(1), 2 (1984)
4 Bliss, R B Dinosaur ABC's (Creation Life El Cajon, California, 1986)

A thyromimetic that decreases plasma cholesterol levels without increasing cardiac activity

A. H. Underwood, J. C. Emmett, D. Ellis, S. B. Flynn, P. D. Leeson, G. M. Benson, R. Novelli, N. J. Pearce & V. P. Shah

Smith Kline and French Research Limited, The Frythe, Welwyn, Hertfordshire AL6 9AR, UK

A number of agents that mimic the ability of the thryoid hormone, T_3, to decrease plasma cholesterol levels are described, one is as effective as T_3 at reducing cholesterol levels and stimulating liver function, but has very little effect on cardiac function and is thus less likely to be toxic The agent may be useful in the treatment of atherosclerosis

CHOLESTEROL concentrations are increased in plasma in hypothyroidism and diminished in hyperthyroidism, treatment of thyroid disease normalizes plasma cholesterol[1] In man[2], and animal models[3 4] with normal thyroid function, relatively low doses of thyroid hormones reduce circulating cholesterol But 3,3′,5-triiodo-L-thyronine (T_3) and L-thyroxine (T_4) cannot be used therapeutically to treat hypercholesterolaemia because of the adverse consequences of some of their other actions, notably on the heart The cardiac effects could either be a direct consequence of the interaction of the hormones with their cardiac receptors, or a result of the characteristic increase in metabolic rate leading, indirectly, to increased cardiac output Several groups of workers have attempted to design analogues that retain hypocholesterolaemic activity but lack cardiac and metabolic effects However, despite claimed selectivity in animals for certain compounds[3,5], all agents that mimic the thyroid hormones (thyromimetics)[6 7], including D-T_4 (ref 7), have shown unacceptable cardiac side effects when tested in man

Thyroid hormones act on the liver in several ways that could increase cholesterol elimination These include induction of low-density lipoprotein (LDL) receptors[8], lecithin-cholesterol acyl transferase[2,9], hepatic lipase[2,10] and lysosomal acid cholesterol esterase[11] Thus an agent that mimics the thyroid hormones in their hepatic but not their cardiac activity could be a valuable agent for treating hypercholesterolaemia Here we describe a new class of selective thyromimetics The most potent of these, SK&F L-94901, after chronic administration, is as active as T_3 on the liver but has about 0 1% of the activity of T_3 on the heart We also show that this compound is a very potent hypocholesterolaemic agent with only a small effect on metabolic rate

Experimental approach

The primary assay used was the induction of the enzyme GPDH (mitochondrial cytochrome *c* 3-phosphoglycerate oxidoreductase) in the heart and liver of a hypothyroid rat Both natural and synthetic thyromimetics caused large (3–6-fold), easily quantified increases in cardiac and hepatic GPDH 48 hours after the administration of a single dose to hypothyroid rats (A H U and D E, unpublished data) The assay gave similar results to ones based on changes in whole body or tissue oxygen consumption, and is thought of as an excellent measure of the effects of thyromimetics on specific organs[12] Secondary screening involved chronic administration, usually oral, of selected compounds to either hypothyroid rats or rats with normal thyroid function, to determine hypocholesterolaemic activity and effects on whole body oxygen consumption (WBOC) and thyroid function

We have used measurements of the affinities of novel compounds for the thyroid hormone receptor of isolated cardiac

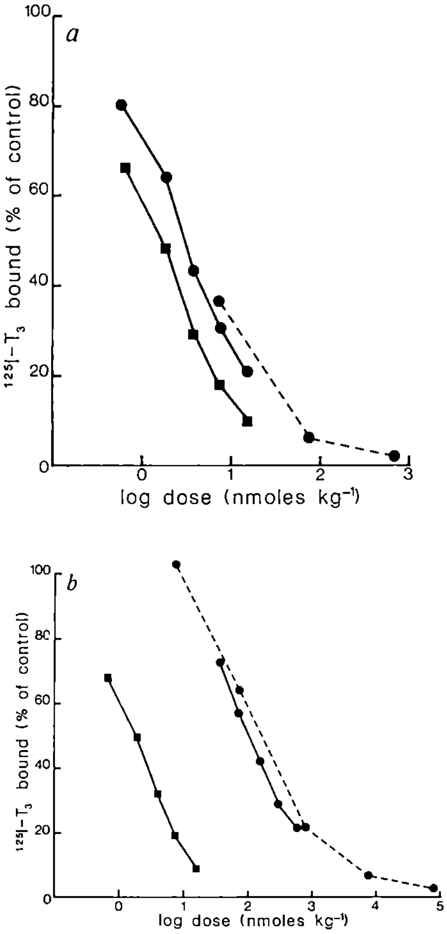
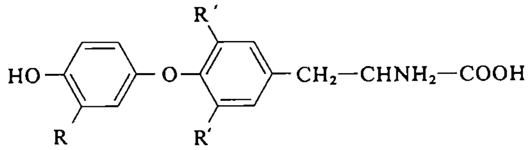
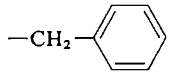
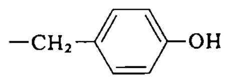
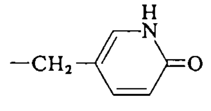
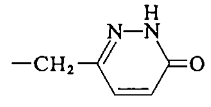
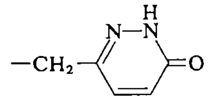


Fig 1 Competition of T_3 and SK&F L-94901 with ^{125}I-T_3 for occupation of nuclear receptors *in vivo* *a*, hepatic nuclei, *b*, cardiac nuclei, ■, unlabelled T_3, ●, SK&F L-94901

Methods Rats were rendered hypothyroid by providing them with drinking water containing 0 3 g l^{-1} of the antithyroid agent 2-mercapto-1-methylimidazole (MTI) for at least 4 weeks They were then given intravenous injections of ^{125}I-T_3 and various doses of unlabelled T_3 or SK&F L-94901 One hour later hepatic and cardiac nuclei were isolated[22] and the ^{125}I and DNA[23] content determined Nuclear ^{125}I was expressed as d p m mg^{-1} DNA and corrected for small variations in the injected dose by dividing by plasma ^{125}I-T_3 determined as that fraction of the plasma ^{125}I content precipitable by trichloracetic acid[22] The experiment with T_3 and the two experiments with SK&F L-94901 (dotted and solid lines) were done on separate occasions The symbols represent the means of three hearts or two livers

Table 1 A comparison of the potencies and receptor binding properties (*in vitro* and *in vivo*) of various substituted thyronines



			Heart			Liver		
				Relative receptor binding		Relative	Relative receptor binding	
SK&F No	R′	R″	Relative potency	*in vitro*	*in vivo*	potency	*in vitro*	*in vivo*
T_3	I	I	100	100	100	100	100	100
L-92904		I	0.5	13	0.3	2.2	15	5.1
L-93236		I	0.1	56	0.7	2.2	17	19
L-93993		I	Inactive	2.7	0.2	1.7	2.3	10
L-94690		I	0.1	4.0	0.4	2.4	2.0	38
L-94901		Br	0.1	2.0	1.3	18	0.9	50

Male, Wistar rats were rendered hypothyroid by administering 0.3 g l^{-1} of the antithyroid agent 2-mercapto-1-methylimidazole (MTI) in drinking water for at least 4 weeks before use. Single doses of thyronines were given intramuscularly in a vehicle of 0.15 M NaCl/0.01 M NaOH (1 ml kg^{-1}) 48 h before the GPGH activities of hearts and livers were determined. A standard '2+2' method was used to determine the potency of SK&F L-92904 and L-93236. Two doses of either T_3, SK&F L-92904 or L-93236 were administered, dose ratios were determined for hearts and livers, and relative potencies calculated. For SK&F L-93993, L-94690 and L-94901, as well as T_3, six doses were administered. For hepatic GPDH the resulting dose response curves were fitted to a hyperbolic curve by a non-linear curve fitting procedure and the mean effective dose (ED_{50}) calculated as described[19]. Potency is defined as the reciprocal of the ED_{50}, relative potency is the ratio of the potency of the novel compound to that of T_3, expressed as a percentage. The ED_{50} for T_3 on cardiac GPDH activity was 0.7×10^{-7} moles kg^{-1}, on hepatic GPDH it was 1.6×10^{-7} moles kg^{-1}. All 6 compounds caused the same maximal increase in hepatic GPDH and are thus full agonists. Only SK&F L-92904 caused the same maximal increase in cardiac GPDH as did T_3. The highest dose used of the remaining compounds was 50 mg kg^{-1} (77–90 μmol kg^{-1}). For SK&K L-93236, L-94690 and L-94901 only this dose significantly increased cardiac GPDH to about 50% of the level achieved with a maximally active dose of T_3. The potencies for these three compounds are, therefore, an approximation based on this figure and the assumption that the compounds would be full agonists if a high enough dose could have been administered. Even at a dose of 50 mg kg^{-1} SK&F L-93993 had no effect on cardiac GPDH. Receptor binding *in vitro* was determined using isolated nuclei. Briefly, nuclei, prepared by sedimentation of homogenates through high-density sucrose solutions, were incubated with ^{125}I-T_3 and increasing concentrations of non-radioactive T_3 or novel compound using published conditions for liver[20] and heart[21]. Bound and free ^{125}I-T_3 were separated by centrifugation and the concentration of non-radioactive ligand reducing the binding of ^{125}I-T_3 by 50% (IC_{50}) was determined. The affinity of T_3 for these receptors (determined from Scatchard plots) was about 10^9 M^{-1} for both heart and liver. *In vivo* receptor binding studies were conducted as described (Fig. 1, legend). Relative receptor binding is defined as the ratio of either the IC_{50} or the ED_{50} for T_3 to that for novel compound expressed as percentage.

and hepatic nuclei[13] to examine the relationship between structure and affinity, and to optimise receptor binding. Indirect evaluations of the capacity of compounds to occupy receptors *in vivo* have also been made using their ability to compete with ^{125}I-T_3 for uptake by nuclei when both ^{125}I-T_3 and the compound are administered simultaneously to rats[14].

Selective thyromimetics

We used the natural hormone, T_3, as the starting point and have confirmed that (1) a diphenylether core structure carrying a 4′-OH and 3,5-halogens is essential for high receptor affinity and (2) 3′-alkyl substituent is an effective replacement for a 3′-iodine atom in T_3 (ref. 15). We then tried to define the optimum 3′-substituent properties for high receptor affinity and hormonal activity. The 3′-benzyl analogue (SK&F L-92904) was made to test the effect of increasing substituent lipophilicity and/or size. This compound had 2.2% of the activity of T_3 in the liver but only 0.5% of the activity of T_3 in the heart. Its receptor affinity is greater, however, being 13% of that of T_3 for isolated cardiac nuclei and 15% for isolated hepatic nuclei (Table 1). This could be a consequence of limited solubility, so the potentially more soluble 4-hydroxybenzyl analogue, SK&F L-93236, was made. This compound showed clear selectivity, being similar in activity to SK&F L-92904 as a hepatic thyromimetic (2.5% of T_3), but much weaker (0.1% of T_3) in the heart (Table 1).

The selectivity of these two compounds is not due to their receptor affinities. The affinities of SK&F L-92904 for the heart and the liver nuclear binding sites are essentially the same, and SK&F L-93236 actually has a higher affinity for cardiac receptors than for those from the liver. Both compounds were less active than would be predicted from their affinities for the receptors *in vitro*. To investigate these anomalies we conducted *in vivo* binding studies with SK&F L-92904 and SK&F L-93236. The results (Table 1) showed that both compounds were much less effective in preventing ^{125}I-T_3 binding to heart nuclei than to liver nuclei. If we asume that the ability of these T_3 analogues to depress nuclear binding of ^{125}I-T_3 *in vivo* is a measure of their capacity to bind to and activate the nuclear thyroid hormone receptor, then the reduced *in vivo* nuclear binding in the heart compared with the liver is broadly consistent with the observed selective thyromimetic activity. We have termed this differential effect on *in vivo* nuclear binding 'selective access'.

A larger group of related compounds was made, the pyridone derivative (SK&F L-93993) was found to be a full agonist in

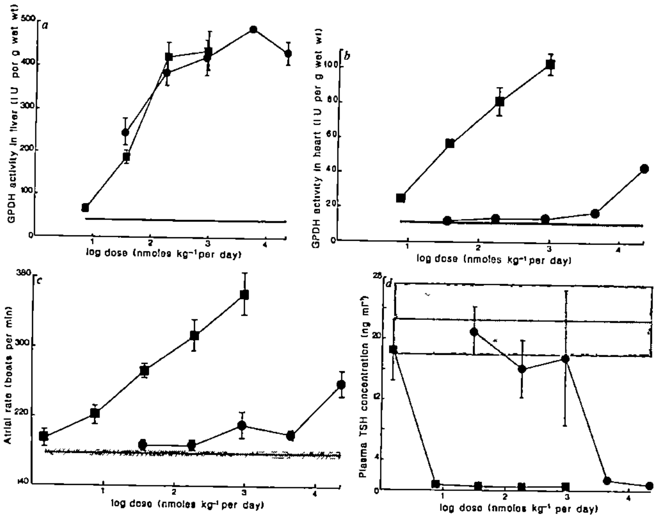


Fig 2 Chronic treatment of hypothryoid rats with T_3 or SK&F L-94901 Effect on *a*, hepatic GPDH activity, *b*, cardiac GPDH activity, *c*, spontaneous rate of beating of isolated right atria, *d*, plasma TSH concentration, ■, T_3, ●, SK&F L-94901

Methods Newly weaned, male, Wistar rats were rendered hypothyroid using MTI as in Fig 1 They were then given seven daily consecutive intramuscular injections of either vehicle (0 15 M NaCl/0 01 M NaOH), or T_3 or SK&F L-94901 solutions in vehicle The rats were anaesthetised with halothane 24 h after the final injection and blood samples withdrawn from the dorsal vena cava Plasma was prepared and stored frozen at −20 °C until TSH could be measured by radio immunoassay (reagents were supplied by the NIH, Bethesda) The right atria were dissected out and mounted in Krebs-Henseleit saline buffered with CO_2/bicarbonate Atrial rates were recorded after a 40-min stabilisation period at 33 °C The remainder of the hearts and samples of liver were frozen on dry ice until their GPDH content could be determined as described[19] Each point represents the mean of between 2 and 7 observations the vertical bars are 2 s e m Horizontal line, mean of values from rats treated with vehicle, cross-hatching, 2 s e m

the liver but inactive in the heart (Table 1) Data from *in vivo* nuclear binding experiments also showed that all selective agonists exhibit selective access Regression analyses suggest a dependence on 3′-substituent polarity and this prompted the synthesis of the more polar pyridazinone (SK&F L-94690), a selective thyromimetic showing a greater degree of selective access But although selective access and selective activity is dependent on 3′-substituent hydrophilicity, the receptor affinity (and hence hormonal activity) was found to be correlated with high 3′-substituent lipophilicity We therefore explored alternative structural modifications to increase activity The 3,5-dibromo analogue (SK&F L-94901) was made because it was likely to be metabolically more stable than T_3 and was found to be highly selective and more potent than the diiodo analogue (Table 1) Like all other selective thyromimetics, SK&F L-94901 shows marked selective access (Table 1, Fig 1)

Note that D-T_4 is not selective as assayed by induction of GPDH, having 7% and 11% of the activity of T_3 in heart and liver respectively

The phenomenon of selective access implies that there is a different relationship between relative receptor binding *in vitro* and *in vivo* for the heart than for the liver The significance of this is discussed below But it is apparent from the data in Table 1 that there is no constant relationship between relative receptor binding *in vivo* and *in vitro*, nor between *in vivo* binding and overall potency for this group of compounds The reasons for these differences have not been examined in detail, but many factors, as well as affinity for isolated receptors, control occupation and hence activation of receptors *in vivo* These factors include binding to plasma proteins, metabolic transformations and the distribution of the thyronines between blood, organs and cellular compartments Finally the possibility that the compounds have differing efficacies cannot be ruled out

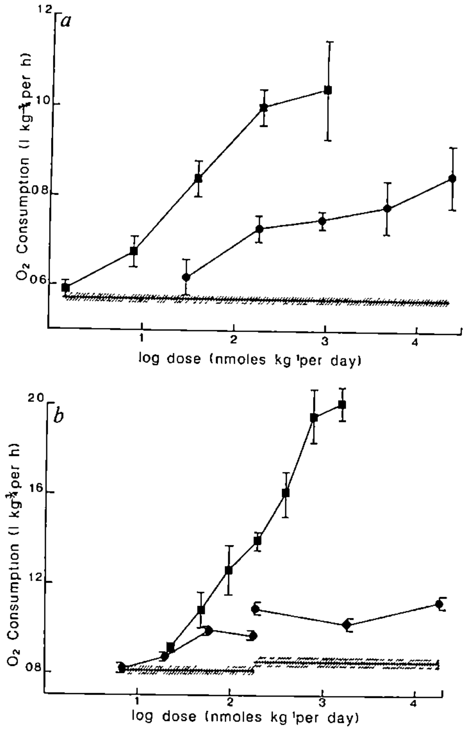


Fig 3 The effects of T_3 and SK&F L-94901 on oxygen consumption *a*, hypothyroid rats, *b*, normal rats, ■, T_3, ●, SK&F L-94901

Methods. Rats were familiarised with the procedure for determining oxygen consumption to ensure that minimum values were obtained, then given seven daily consecutive doses of vehicle (0 15 M NaCl/0 01 M NaOH), T_3 or SK&F L-94901 Oxygen consumption was measured 24 h after the final dose and expressed as litres of O_2 (at Standard temperature and pressure (STP)) consumed $kg^{-3/4}$ per h since this is independent of body weight[24] Symbols represent the means of between 3 and 7 rats, vertical bars, 2 s e m , horizontal lines, mean oxygen consumption in vehicle-treated controls ($n = 12$ in *a* and 5 in *b*) *a*, hypothyroid rats, as in Fig 2, given intramuscular injections at 1 ml kg^{-1} Oxygen consumption, at 30 °C, was determined by a modification of the 'flow-through' technique[26] and was significantly greater than in controls for all groups except those given the lowest doses of either compound *b*, normal rats dosed orally Oxygen consumption was measured, at 29 °C, as described[19] Three separate experiments over different dose ranges are shown on this graph The mean oxygen consumption in the controls ($n = 5$) for the two experiments with SK&F L-94901 ate shown as horizontal lines below the relevant curves hatching 2 s e m The mean oxygen consumption of vehicle-treated controls for the T_3 dose-response curve was intermediate at 0 84 ± 0 02 (4) l O_2 at STP $kg^{-3/4}$ per h Oxygen consumption was significantly greater than control in all treated groups except for those given the two lowest doses SK&F L-94901 and three lowest of T_3

The most active selective compound

After a single dose SK&F L-94901 has about 20% of the potency of T_3 in the liver and 0 1% in the heart (Table 1) Chronic treatment enhances the hepatic potency of SK&F L-94901 After seven daily intramuscular doses SK&F L-94901 is as potent as T_3 as a thyromimetic in the liver of hypothyroid rats, but still only had 0 1% of the activity of T_3 on cardiac GPDH activity and on the spontaneous rate of beating of isolated atria (Fig 2*a*, *b* and *c*) In rats with normal thyroid function, given SK&F L-94901 or T_3 orally for seven days, a very similar result was obtained The two compounds were equipotent in the liver but only T_3 had an effect on cardiac GPDH SK&F L-94901 was inactive at a dose 200 times greater than the cardiac ED_{50} for T_3 In this experiment the effects of SK&F L-94901 and T_3 on

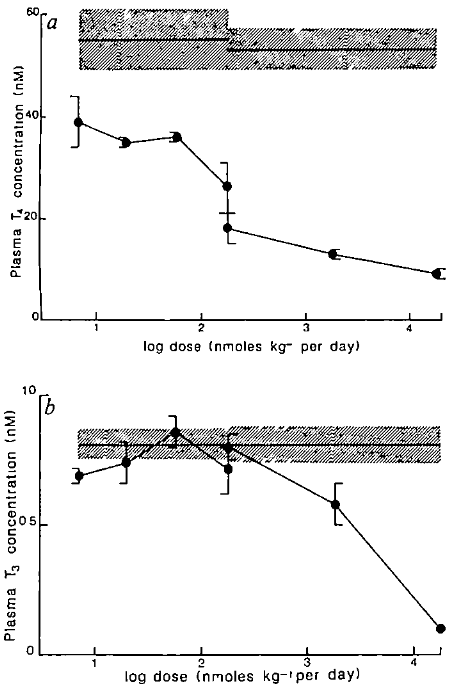


Fig 4 Effect of SK&F L-94901 on *a*, plasma T_3 and *b*, T_4 concentrations Male euthyroid rats were given seven daily sonsecutive oral doses of KS&F L-94901 and samples of blood from the dorsal aorta collected 24 h after the final dose under halothane anaesthesia Plasma T_3 and T_4 concentrations were determined by radioimmunoassay (RIA) using commerically available kits (Amerlex T_3 and T_4 RIA kits, Amersham) which are validated for use with rat plasma Values shown are the means of 4 or 5 rats, vertical bars, 2 s e m , horizontal lines, vehicle treated controls, cross-hatching, 2×s e m of controls Two separate experiments are shown, the control values for each experiment are shown above *a*, T_3 concentrations Only the highest dose of SK&F L-94901 had a statistically significant effect and this was to reduce T_3 concentrations to below the limit of detectability (0 15 nM) *b*, T_4 concentrations Mean values in all treated groups, except those given the lowest dose of SK&F L-94901, were significantly smaller than in vehicle-treated controls

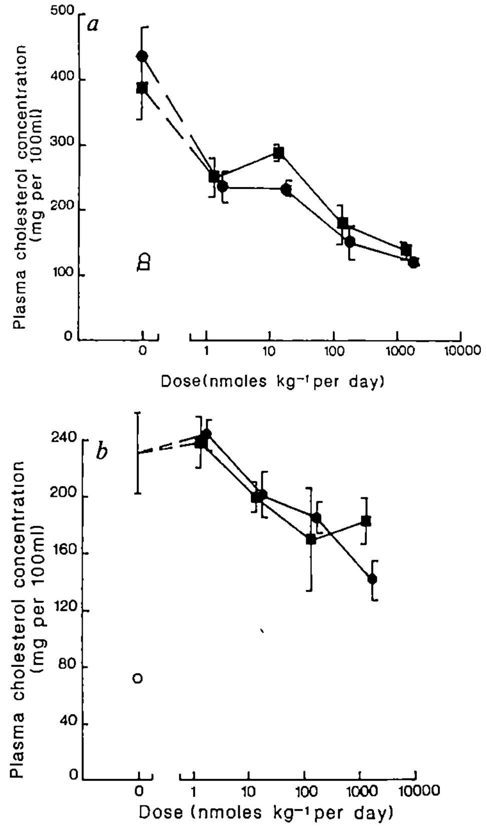


Fig 5 Effects of T_3 and SK&F L-94901 on plasma cholesterol concentrations *a*, Hypothyroid rats, fed with cholesterol *b*, normal rats fed with cholesterol, ■, T_3, ●, SK&F L-94901, open symbols indicate groups fed with a normal diet Male rats were rendered hypercholesterolaemic by feeding a diet containing 1 5% cholesterol and 0 5% cholic acid for 3 weeks Seven daily consecutive oral doses (5 ml kg^{-1}) of vehicle (0 15 M NaCl/0 01 M NaOH), T_3 or SK&F L-94901 were given Groups of male rats receiving a normal diet were also dosed with vehicle for seven days Blood was taken under anaesthesia 24 h after the final dose and total plasma cholesterol measured with commercially available kit (Diagnostica Merck, E Merck) using cholesterol oxidase The symbols represent the mean values from groups of 4–5 rats, the vertical bars are the s e m

renal GPDH were also measured, both compounds were fully active and roughly equipotent

Chronic administration of either SKF L-94901 or T_3 increases the minimal WBOC of rats, although the character and magnitude of the responses are very different In hypothyroid rats the highest dose of T_3 used roughly double the WBOC SK&F L-94901 increases the WBOC by 50% at its highest dose, with a shallower dose-response curve (Fig 3*a*) This phenomenon was even more apparent in rats with normal thyroids, where the maximum response to SK&F L-94901 was about one quarter of that to T_3, and the dose-response curve to SK&F L-94901 was esentially flat over a considerable dose-range (Fig 3*b*) This would be expected if a smaller number of organs gave a thermogenic response to SK&F L-94901 than to T_3 administration

Both SK&F L-94901 and T_3 diminish plasma thyroid stimulating hormone (TSH) concentrations in hypothyroid rats after chronic treatment, but have dramatically different potencies The lowest dose of SK&F L-94901 that reduced plasma TSH significantly was more than 1,000 times greater than the lowest effective dose of T_3 (Fig 2*d*) When administered chronically to rats with normal thyroids only the highest dose of SK&F L-94901 reduces plasma T_3 concentrations significantly (Fig 4*a*) This is similar to the lowest dose that significantly reduced plasma TSH in hypothyroid rats, implying that the effect of SK&F L-94901 on plasma T_3 could have been mediated by an action on pituitary thyrotrophs By contrast, SK&F L-94901 reduces circulating T_4 significantly in a dose-related manner, although no dose eliminated T_4 altogether (Fig 4*b*) This maintenance of plasma T_3 while T_4 concentrations decrease is unusual and could be due to increased conversion of T_4 to T_3 or diminished elimination of T_3

Unlike man, normal rats have very low circulating plasma cholesterol concentrations (about 50 mg per 100 ml), and almost all of this is in the high-density lipoprotein fraction (HDL) Thyroid hormones do not reduce total plasma cholesterol in this model Hypocholesterolaemic activity was therefore determined in rats with artificially elevated plasma cholesterol produced by feeding a cholesterol-rich diet Ultracentrifugation of plasma from cholesterol-fed rats revealed a discrete peak of cholesterol in the HDL fraction and a more diffuse fraction of lower density (reported to contain βVLDL and LDL[16]) When hypercholesterolaemic hypothyroid rats were given SK&F L-94901 orally for 7 days the concentration of cholesterol in their plasma was reduced in a dose-related fashion This was mainly due to a reduction in the lower density fraction leading to an increase in the proportion of cholesterol carried on HDL A daily dose as low as 1 μg per kg per day caused a significant reduction compared to a vehicle-treated control group Higher doses reduced plasma cholesterol concentrations to those found in animals given a normal diet T_3 was also active under these conditions and the two compounds were of similar potency (Fig 5*a*) Rats with normal thyroids are less sensitive to both

T_3 and SK&F L-94901, but even under these circumstances, the two compounds have similar hypocholesterolaemic potency (Fig 5*b*)

Discussion

Unlike many hormone agonists and antagonists, the organ selectivity of SK&F L-94901 does not appear to result from differences in the receptors in the relevant tissues, as cardiac and hepatic nuclear receptor affinities are very similar for several hundred compounds (unpublished data), strongly suggesting that the two receptors are identical The evidence that interaction with the nuclear T_3-binding protein is an obligatory step in the action of these hormones is now very strong, but factors controlling the access of the hormone to the nuclear proteins may exert a strong influence on activity Indeed Oppenheimer and Schwartz[17] have recently shown that both the plasma membrane and the nuclear membrane do have such a role In particular, their experiments suggested that the free concentration of T_3 inside the nucleus is much greater than that in the cytoplasm, and they propose that some sort of concentrating mechanism or 'nuclear pump' must be in operation This nuclear pump could be the source of selective access if the liver was more active *in vivo* in pumping SK&F L-94901 into the nucleus than was the heart A corollary of this would be that the nuclear pump does not operate in our *in vitro* binding assay Alternatively, hepatic cytosol may contain higher concentrations of SK&F L-94901 than does cardiac cytosol—either because it is actively accumulated by hepatocytes, or excluded by cardiac myocytes

The similar potencies of SK&F L-94901 as an hepatic thyromimetic and as an hypocholesterolaemic agent support the idea that the actions of thyroid hormones on the liver are, at least in part, responsible for their ability to reduce plasma cholesterol concentrations This is corroborated by the finding that in cats with normal thyroid function and a genetically hypercholesterolaemic strain of rabbit[18] heterozygous for an LDL-receptor deficiency, SK&F L-94901 can also reduce circulating plasma cholesterol levels by about 30% Again, LDL-cholesterol was preferentially reduced, resulting in an increased HDL LDL cholesterol ratio Humans with normally functioning thyroid are relatively sensitive to the hypocholesterolaemic effects of T_3, doses of a few microgrammes per kilogramme of body weight having a significant effect[2] These doses are similar to those reducing plasma cholesterol in the hypercholesterolaemic hypothyroid rat and suggest that SK&F L-94901 could reduce plasma cholesterol in man (and perhaps prevent atherogenesis) at doses far below those having a direct cardiac effect, and which would have minimal or no activity on basal metabolic rate and the thyroid-pituitary axis The necessary toxicological evaluations are under way to allow the administration of SK&F L-94901 to man

We thank Professor Sir James Black, Professor C R Ganellin and Dr R Pitt-Rivers, for support, advice and helpful discussions, G R Showell and H D Prain for contributions to the design and synthesis of compounds, Dr P C Sivelle for measurements of hormones, and I Busolini, M Knowles and P Nixon for experimental support

1 Aviram M Luboshitzky R & Brook J G *Clin Biochem* **15**, 62-66 (1982)
2 Hansson, P, Valdemarsson, S & Nilsson Ehle P *Horm. Metab Res* **15**, 449-452 (1983)
3 Boyd, G S & Oliver, M F *J Endocrinol* **21**, 25-32 (1960)
4 Engleken, S F & Eaton P R *Atherosclerosis* **38**, 177-188 (1981)
5 Cuthbertson, W F J, Flcoate, P V, Ireland, D H, Mills, D C B & Shearley, P *J Endocrin* **21**, 45-63 (1960)
6 Boyd, G S & Oliver, M F *J Endocrin* **20**, 33-43 (1960)
7 The Coronary Drug Project Research Group *J Am med Ass* **220**, 996-1008 (1972)
8 Scarabottolo, L, Trezzi F, Roma, P & Catapano, A L *Atherosclerosis*, **59**, 329-333 (1986)
9 Ridgway, N D & Dolphin, P J *J Lipid Res* **26**, 1300-1313 (1985)
10 Hulsman, W C, Oerlemans, M C & Geelhoed Mieras, M M *Biochem biophys Res Comm* **79**, 784-788 (1977)
11 Severson D L & Fletcher, T *Biochim biophys Acta* **675**, 256-264 (1981)
12 Hoffman, W W, Richert, D A & Westerfeld, W W *Endocrinology* **78**, 1189-1197 (1966)
13 Oppenheimer J H, Schwartz, H L, Surks, M I, Koerner, D & Dillman, W H *Recent Prog in Hormone Res* **32**, 529-565 (1976)
14 Oppenhiemer, J H, Koerner, D, Schwartz, H L & Surks, M I *J clin Endocr Metab* **35**, 330-333 (1972)
15 Weisgraber, K H & Mahley, R W in *Lipoprotein Studies of Non human Species* (eds Lewis, L A & Oppet, J J) 13-103 CRC Handbook of Electrophoresis, vol IV (CRC Press, Baton Rouge 1983)
16 Jorgensen, E C in *Hormonal Proteins and Peptides*, Vol VI Thyroid Hormones (ed Li, C H) 108-204 (Academic, New York, 1978)
17 Oppenheimer, J H & Schwartz, H L *J clin Invest* **75**, 147-154 (1985)
18 Richards, T, Horlock, H & Gallagher, P *Animal Technology* **37**, 1-6 (1986)
20 Benson M G *et al Biochem Pharmacol* **33**, 3143-3149 (1984)
21 Samuels, H H & Tsai, J S *Proc natn Acad Sci USA* **70**, 3488-3492 (1973)
22 Koerner, D, Schwartz, H L, Surks, M I & Oppenheimer, J H *J biol Chem* **250**, 6417-6423 (1975)
23 Oppenheimer, J H, Schwartz, H L, Dillman, W L & Surks, M I *Biochem biophys Res Commun* **55**, 544-550 (1973)
23 Burton, K *Biochem. J* **62**, 315-323 (1956)
24 McMahon, T *Science* **179**, 1201-1205 (1973)
25 Kappagoda, C T & Linden R J *Cardiovasc Res* **6**, 589-598 (1972)

The *Tetrahymena* ribozyme acts like an RNA restriction endonuclease

Arthur J. Zaug, Michael D. Been & Thomas R. Cech

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA

A shortened form of the Tetrahymena *self-splicing ribosomal RNA intervening sequence acts as an endoribonuclease, catalysing the cleavage of large RNA molecules by a mechanism involving guanosine transfer The sequence specificity approaches that of the DNA restriction endonucleases Site-specific mutagenesis of the enzyme active site alters the substrate sequence specificity in a predictable manner, so that endoribonucleases can be synthesized to cut at a variety of tetranucleotide sequences*

THE *Tetrahymena* ribosomal RNA intervening sequence (IVS) is a catalytic RNA molecule or ribozyme It mediates RNA self-splicing, excising itself from the large ribosomal RNA precursor, and then converts itself to a circular form[1,2] In these reactions, the splice sites and circularization sites can be viewed as intramolecular substrates for an enzymatic activity of the IVS RNA[3]

This view is supported by studies of the L-19 IVS RNA, a linear form of the IVS that lacks the first 19 nucleotides Because it has no circularization sites, the L-19 IVS RNA cannot undergo intramolecular reactions[3], but it retains activity and can catalyse cleavage-ligation reactions on other RNA molecules[4] When provided with oligo(cytidylic acid) as a substrate, the L-19 IVS RNA acts as an enzyme with nucleotidyltransferase (poly(C) polymerase) and phosphodiesterase (ribonuclease) activities[4] With 3′-phosphorylated oligo(C) substrates, the same ribozyme

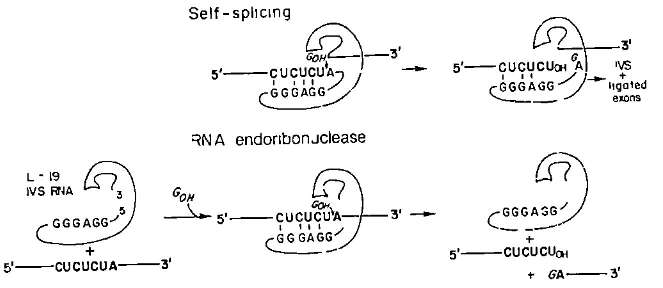


Fig 1 A model for the endonuclease action of L-19 IVS RNA The mechanism is an intermolecular version of the first step of pre-rRNA self-splicing Thin letters and lines, IVS sequences, bold letters and thick lines, exon sequences (top) or substrate RNA sequences (bottom), the G in italics is a free guanosine nucleotide or nucleoside

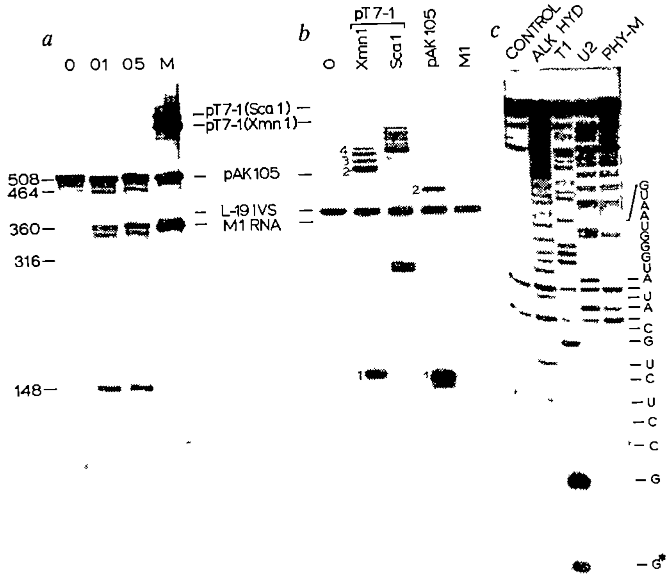


Fig 2 The L-19 IVS$_\beta$ RNA cleaves large RNA substrates with transfer of guanosine *a*, Uniformly labelled 0 6 µM pAK105 RNA (508 nt) incubated with 0 2 µM L-19 IVS$_\beta$ RNA and 0, 0 1 or 0 5 mM GTP (unlabelled) for 1 h under conditions described below M, Mixture of 4 substrate RNAs as molecular weight markers *b*, Various ^{3}H-RNA substrates (1 µg each) incubated with 0 2 µM L-19 IVS$_\beta$ RNA for 1 h, under the same reaction conditions as *a*, but with 120 µM [α-^{32}P]GTP The autoradiogram reveals [^{32}P]GTP-labelled products only The L-19 IVS$_\beta$ RNA also becomes ^{32}P-labelled during incubation *c*, Nucleotide sequence of the cleavage product pAK105(1) determined by the enzymatic method[33] Some nucleotides could not be assigned due to the presence of a band in the untreated (control) RNA sample G*, labelled GTP joined to the RNA during the reaction

Methods L-19 IVS RNA was made by methods similar to those described[4], except that a different RNA polymerase-template system was used Plasmid pT7-TT1A3 (which contains a T7 RNA polymerase promoter, a 42-bp 5' exon, the entire 413-bp IVS, and an 82-bp 3' exon) was cleaved with *Eco*RI and transcribed with bacteriophage T7 RNA polymerase[34] The transcripts were incubated further under self-splicing, circularization, and site-specific hydrolysis conditions to produce L-19 IVS RNA[3,4] The 3'-terminal guanosine was then removed by periodate oxidation and β-elimination[12] to yield L-19 IVS$_\beta$ RNA Substrate RNAs were produced by T7 RNA polymerase transcription of *Bam*HI-cleaved pAK105 (ref 13), *Xmn*I- or *Sca*I-cleaved pT7-1 (purchased from U S Biochemical Corp), and *Sna*BI-cleaved pDW27, which encodes M1 RNA (obtained from D Wahl and N Pace) Substrate RNAs were incubated with L-19 IVS$_\beta$ RNA in 5 mM MgCl$_2$, 10 mM NaCl, 50 mM Tris-HCl, *p*H 7 5 at 50 °C, GTP was present at the concentration indicated Reactions were stopped by adding EDTA to a final concentration of 25 mM Products were analysed by electrophoresis in 4% polyacrylamide, 8 M urea gels and subjected to fluorography (*a*) or autoradiography (*b*)

acts as a phosphotransferase and an acid phosphatase[5] A key feature of all four of these reactions is the formation of a covalent enzyme-substrate intermediate in which a nucleotide or phosphate is esterified through the 3'-O of G414, the 3' terminal guanosine of the IVS RNA

Here we describe a fifth enzymatic activity of the *Tetrahymena* ribozyme It cleaves RNA molecules at sequences resembling the 5' splice site of the rRNA precursor A free guanosine nucleotide is added to the 5' end of the downstream RNA fragment during cleavage This allows one product to be end-labelled during the reaction The reaction is analogous to the first step of pre-rRNA self-splicing (Fig 1) Cleavage does not require nucleotide G414 of the ribozyme, thus unlike the first four activities, it does not involve formation of a covalent enzyme-substrate intermediate

The ribozyme cleaves very specifically after the sequence CUCU, under stringent conditions it discriminates against sites that match three out of four of these nucleotides This specificity approaches that of the DNA restriction endonucleases[6] Site-specific mutations in the active site of the IVS RNA, the internal guide sequence[7,8] or 5' exon-binding site[9-11], alter the sequence specificity of the ribozyme in a predictable manner This will allow us to develop a set of sequence-specific RNA endonucleases that may be useful as tools for studying the molecular biology of RNA

Sequence-specific cleavage of large RNAs

The L-19 IVS RNA enzyme was prepared by incubation of pre-rRNA under conditions that promote self-splicing, circularization, and site-specific hydrolysis[3,4] The 3'-terminal guanosine (G414) was then removed from the L-19 IVS RNA by periodate oxidation followed by β-elimination[12] As expected, the resulting ribozyme (L-19 IVS$_\beta$) has greatly reduced activity as a nucleotidyltransferase, assayed using [^{32}P]-p(C)$_5$ as a substrate (data not shown) But when GTP is added the ribozyme can cleave p(C)$_5$ and large RNA molecules For example, the 504-nucleotide (nt) pAK105 transcript[13], a fragment of mouse β-globin pre-mRNA containing the first intron, is cleaved to give major fragments of 148, 360 and 464 nt, and some minor fragments (Fig 2*a*) As shown below, the 360- and 148-nt fragments can be explained as the 5' and 3' products of cleavage at position 360 The 464-nt fragment is the 3' product of cleavage at position 44, the 5' 44-nt fragment being too small to be observed The fact that little 316-nt RNA (the expected product of cleavage at both position 44 and 360) is observed indicates partial digestion, with few molecules cleaved more than once

Cleavage requires Mg^{2+} (optimum at 10-20 mM MgCl$_2$) and is essentially independent of monovalent cation in the range 0-200 mM NaCl The *p*H optimum is between 7 5-8 0, and the temperature optimum is ~50 °C Although β-eliminated L-19 IVS RNA catalyses the cleavage reaction, the enzyme works at the same rate whether or not G414 has been removed (data not shown) We explain this by supposing that, at saturating GTP concentrations, the attack by GTP on the substrate competes very effectively with attack by G414

When [α-^{32}P]GTP is included in the reaction of pAK105 RNA, the 148- and 464-nt cleavage products are labelled (Fig 2*b*) Direct sequencing of these labelled RNA fragments (Fig 2*c*) showed that cleavage and GTP-addition occur at nucleotides 44 and 360 in the sequence, such that the downstream cleavage products are 5'-GTP labelled The bonds formed by GTP addition are sensitive to RNase T$_1$, confirming that the GTP is covalently added through its 3'-O by a normal 3'-5' phosphodiester bond Reaction of the 841-nt pT7-1 (*Xmn*I) RNA (essentially pBR322 sequences) produced four major labelled fragments These were sequenced in the same manner The pT7-1 RNA with an extra 122 nt at its 3' end was produced by transcription of pT7-1 DNA that had been cleaved at the *Sca*I site, the

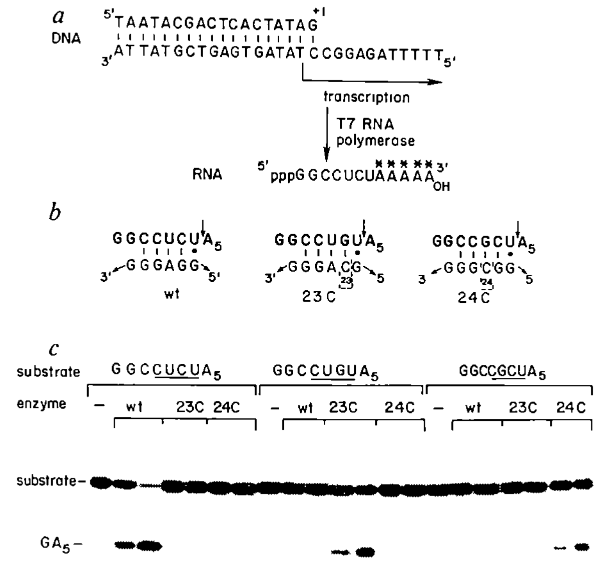


Fig 3 Three different ribozymes can distinguish between sequences that differ by a single base change in the recognition element *a*, Synthesis of defined oligoribonucleotide substrates by the method of Lowary *et al*[19] DNA was transcribed with T7 RNA polymerase in the presence of [α-^{32}P]ATP to produce oligoribonucleotides labelled with ^{32}P in the positions indicated (*) b, Proposed interactions between the three oligoribonucleotide substrates (top strands, bold) and the active sites of the matched ribozymes (bottom strands) Arrows, sites of cleavage and guanosine addition *c*, Cleavage of three oligoribonucleotide substrates by wild-type and variant L-19 IVS RNAs, assayed by 20% polyacrylamide, 7 M urea gel electrophoresis (–), untreated substrate Lanes are run in pairs, 1 0 μM ^{32}P-labelled substrate was incubated with 0 125 μM ribozyme for 15 min (left lane) or 60 min (right lane) at 50 °C in 10 mM MgCl$_2$, 10 mM NaCl, 50 mM Tris-HCl, *p*H 7 5, 0 5 mM GTP (unlabelled), 2 5 M urea Because the ^{32}P is confined to the nucleotides downstream from the cleavage site, only the smaller product is apparent The identity of the GA$_5$ product was confirmed by treatment of the substrate and reaction mixtures with RNase T$_2$ and monitoring the transfer of ^{32}P to Gp A band migrating above GA$_5$ in lanes 2 and 3 appeared in some experiments but has not been identified

Methods Substrates were prepared by transcription of deoxyoligonucleotides synthesized on an Applied Biosystems DNA Synthesizer The same promoter top strand was used in each transcription The bottom strand, which contains promoter and template sequences, was varied to obtain the different RNA substrates The DNA was transcribed with purified phage T7 RNA polymerase[34] as described[19] Variant ribozymes are described by Been and Cech[11] The L-19/23C ribozyme was prepared by transcription of pBG/-2G 23C and treatment of RNA under site-specific hydrolysis conditions[11] The 24C ribozyme was similarly prepared from transcripts of pBG/–3G 24C The variant ribozymes were not subjected to β-elimination to remove their 3'-terminal G

Table 1 Sites of cleavage of large RNA substrates by the wild-type L-19 IVS$_\beta$ RNA

Substrate	Site	Size (nt)*	Sequence†
			–6 –1
pAK-105 (β-globin pre-mRNA)	1	148	UCCUCU↓GCCUC
	2	464	AACUCU AAGAG
pT7-1 (pBR322)	1	145	UCCCUU UUUUG
	2	556	CACUCU CAGUA
	3	603	UAUUUU CUCCU
	4	805	UCCUCU AGAGU
Consensus			
observed			$\mathrm{U}^{\mathrm{C}}_{\mathrm{A}}\mathrm{CUCU}\downarrow\mathrm{N}$
expected‡			$\mathrm{C}^{\mathrm{C}}_{\mathrm{U}}\mathrm{CUCU}\downarrow\mathrm{N}$

* Size of the GTP-labelled fragment

† Sequences are listed in 5'→3' direction Arrows indicate sites of cleavage and guanosine addition

‡ The expected consensus sequence is based on the known sequence at the end of the 5' exon (CUCUCU) modified by the possibility that either a C or a U at position –5 might be able to pair with the G in the active site At position –1, C may work less well than U[7,27,35]

labelled products increased in molecular weight as expected In all reactions, including those in which substrate RNA was omitted, the L-19 IVS$_\beta$ RNA became labelled, perhaps by guanosine exchange[14,15] The sites of self-labelling were heterogeneous

The sequence near the 5' end of each end-labelled product was compared with the known sequence of the RNA to identify the nucleotides preceding the site of cleavage (Table 1) All the cleavage sites are preceded by four pyrimidines, the consensus sequence is CUCU This is exactly the tetranucleotide sequence expected to be an optimal target for the ribozyme The importance of nucleotides at –5 and –6 from the cleavage site is not clear, but the absence of G residues may be significant There is no apparent sequence preference downstream from the cleavage site, all four nucleotides are found at position +1

The potential sites in the RNA that were not cleaved by the ribozyme also provide information about the specificity For the pAK105 RNA, only one CUCU site was not cleaved (Cleavage at this site would produce a labelled 378-nt RNA) Many sites that match CUCU in three out of four positions were not cleaved These include 17 CUMU sequences (where M ≠ C) and 7 CNCU sequences (where N ≠ U) In pT7-1 RNA, cleavage was observed at both of the CUCU sequences in the RNA, but at only one of the 15 UUUU sequences present Thus, the ribozyme has a strong preference for cleavage after the tetranucleotide CUCU

M1 RNA, the RNA subunit of *Escherichia coli* RNase P[16], was not cleaved by L-19 IVS$_\beta$ RNA under standard reaction conditions (Fig 2*b*) M1 RNA contains the sequence UCCUCU, which should be an excellent target site However, this sequence is part of a stable hairpin stem[17,18], which presumably makes it unavailable as a substrate Denaturation of M1 RNA with glyoxal allows this site to be cleaved by L-19 IVS$_\beta$ RNA (unpublished data), in support of the interpretation that RNA secondary structure can inhibit cleavage

Active-site mutations alter substrate specificity

We then studied the substrate specificity in a defined system where the substrate RNA had no secondary structure that might affect the cleavage reaction Oligoribonucleotide substrates were made using the phage T7 RNA polymerase transcription method (Fig 3*a*) One substrate contained a perfect match to the tetranucleotide consensus sequence Two other substrates had single-base changes giving a three-out-of-four match to the consensus

These substrates were tested with the wild-type L-19 IVS$_\beta$ RNA and with two altered ribozymes The two variants have single-base changes in the 5' exon-binding site that alter the sequence specificity of the first step in pre-rRNA self-splicing[11] The 23C variant (G→C at position 23 of the L IVS RNA) is expected to recognize CUGU substrates, and the 24C (A→C at position 24) variant should recognize CGCU (Fig 3*b*) In the presence of 2 5 M urea, each substrate is cleaved exclusively by the ribozyme that can make a perfectly base-paired enzyme-substrate complex (Fig 3*b*, *c*) We find that 2 5 M urea or 15% formamide greatly increases specificity, allowing each ribozyme to differentiate between substrates with a single base change in the recognition sequence Our current model is that these denaturants destabilize the base-pairing between the substrate and the active site nucleotides of the ribozyme, thereby discriminating against mismatched complexes

When variant ribozymes were incubated with the 504-nt pAK105 RNA, each ribozyme gave a different pattern of cleavage products (data not shown) One major cleavage site was mapped for three variant ribozymes, including a 25C variant (G→C at position 25) The sites cleaved by the 23C, 24C and

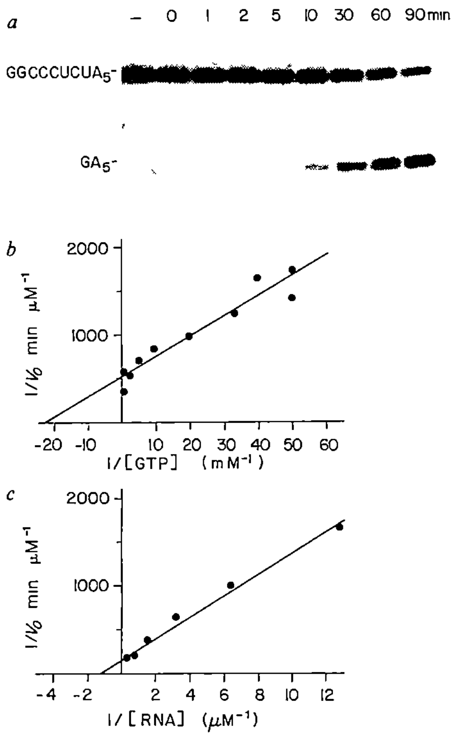


Fig. 4 Kinetic analysis of the RNA endoribonuclease reaction. *a*, The oligoribonucleotide substrate (2.5 µM) was incubated with wild-type L-19 IVS$_\beta$ RNA (0.2 µM) as in Fig. 3, except that urea was omitted. *b*, Kinetics of cleavage of GGCCCUCUA$_5$ as a function of GTP concentration. RNA substrate concentration was kept constant at 2.5 µM. *c*, Kinetics of cleavage as a function of RNA substrate concentration, with GTP concentration kept constant at 0.5 mM.

Methods. Products were separated by polyacrylamide gel electrophoresis. Each gel was cut into strips using the autoradiogram as a guide and the radioactivity in the unreacted substrate and the GA$_5$ product was determined by liquid scintillation counting. The initial rate of cleavage (v_0) was determined from a semilogarithmic plot of the fraction of reaction as a function of time. The graphs are linear least-squares fits to a plot of $1/v_0$ against inverse substrate concentration.

25C ribozymes are preceded by CCCUGU, UCUGCU, and CUGUCU, respectively; the underlining indicates the base that forms a G·C base pair with the mutated nucleotide in the variant ribozyme. Each of these sequences can form 6 continuous base-pairs with the active site nucleotides of the variant ribozyme.

Cleavage is catalytic

The time course of cleavage of the oligonucleotide substrate GGCCCUCU↓AAAAA (the arrow designates the cleavage site) by wild-type ribozyme is shown in Fig. 4*a*. The reaction of 2.5 µM substrate with 0.2 µM ribozyme is 66% complete in 90 minutes. Thus, it is readily apparent that the ribozyme is acting catalytically.

The reaction rate was determined at a series of substrate concentrations. The kinetics are adequately described by the Michaelis–Menten rate law. The dependence of the rate on GTP concentration is shown in the form of a Lineweaver–Burk plot in Fig. 4*b*. The K_m for GTP is 44 µM. The dependence of the rate on RNA substrate concentration at saturating GTP concentration is shown in Fig. 4*c*. The K_m for this oligoribonucleotide substrate is 0.8 µM, and k_{cat} is 0.13 min^{-1}. Thus under V_{max} conditions the enzyme turns over about 8 times per hour.

Discussion

We find that the L-19 IVS RNA acts like an RNA restriction endonuclease. It catalyses the sequence-specific cleavage of other RNA molecules while transferring guanosine to the downstream RNA fragment. The enzyme has a strong preference for cleaving RNA after the sequence CUCU (Table 1), and in a solution containing 2.5 M urea it cleaves after CUCU, but ignores the related sequences CUGU and CGCU (Fig. 3*c*). Additional sequences must be tested to provide a complete picture of the sequence specificity of the ribozyme, and to examine the effect of the nucleotides at positions −5 and −6 on the efficiency of cleavage. Its specificity may approach that of the DNA restriction endonucleases, most of which recognize a particular 4 or 6 base-pair sequence.

The endoribonuclease reaction is analogous to the first step of pre-rRNA self-splicing (Fig. 1). Both the enzymatic and the self-splicing reactions use the same two binding sites, an oligopyrimidine-binding site and a guanosine-binding site, to achieve specificity and to contribute to catalysis. The oligopyrimidine-binding site is also known as the 5′ guide sequence[8] or the 5′ exon-binding site[9-11]. Its role in self-splicing has been defined by analysis of single-base mutations and second-site suppressor mutations[8,11,20]. The same nucleotides bind the substrate in the endoribonuclease reaction, as shown by the change in substrate specificity of the mutant enzymes (Fig. 3). The altered specificity follows the rules of Watson–Crick base pairing. The general features of the guanosine-binding site involved in self-splicing have been described[21,22], but the site has not been localized to a particular set of nucleotides. The endoribonuclease activity appears to use the same guanosine-binding site, as in both cases guanosine is as active as GTP, whereas UTP, CTP, ATP and dGTP have little if any activity (data not shown). In addition, the K_m of 44 µM for GTP observed in this study agrees reasonably well with the value of 32 ± 8 µM determined for self-splicing under somewhat different reaction conditions[22].

The endoribonuclease activity of L-19 IVS RNA does not require the 3′-terminal guanosine (G414). In this respect it differs from the nucleotidyl transfer, phospho transfer and hydrolytic activities, in which a covalent enzyme-substrate complex, apparently necessary for reaction, is formed through G414[4,5]. Thus the L-19 IVS RNA is not restricted to reaction mechanisms involving formation of a covalent enzyme-substrate intermediate. It can also catalyse bisubstrate reactions by a single-displacement mechanism. Ribonuclease P, an RNA enzyme that catalyses cleavage by hydrolysis rather than by transesterification, also appears to act without formation of a covalent intermediate[23,24].

The L-19 IVS RNA endoribonuclease activity reported here appears to require single-stranded RNA substrates. Work recently reported by Szostak[25] suggests that a smaller version of the *Tetrahymena* IVS RNA, lacking the 5′ exon-binding site, may have an endoribonuclease activity that requires a base-paired substrate. The substrate tested[25] was an RNA fragment containing the end of the 5′ exon paired with the 5′ exon-binding site. However, this RNA 'substrate' included a substantial portion of the IVS RNA, so it is unclear whether the ribozyme has endoribonuclease activity with double-stranded RNA substrates in general.

The enzymatic cleavage reaction has some implications for *Tetrahymena* pre-RNA self-splicing. The observed sequence-specificity of the enzymatic reaction provides further evidence that the CUCU sequence immediately preceding the 5′ splice site is both necessary and sufficient to designate the site. RNA structure models[7,26,27] and data from the *trans*-splicing reaction[9,10] and genetic studies[8,11,15] also support this conclusion. Now that the first step of self-splicing can be mimicked using a short oligonucleotide substrate, analysis of this step should become easier. For example, the uridine nucleotide immediately preceding the site of guanosine addition can be replaced with modified uridine This uridine is conserved in Group I intervening sequences[7,27], modifications would allow detailed investigation of the interactions involving this uridine

Sequence-specific endoribonucleases may be useful in the study of RNA, in ways analogous to DNA restriction endonucleases[6] The pattern of restriction fragments could be used to establish sequence relationships between two related RNAs, and large RNAs could be cleaved to fragments of a size more useful for study The 4-nucleotide specificity of the ribozyme is ideal for cleavage of RNAs of unknown sequence, an RNA of random sequence would have an average of 1 cleavage site every 256 bases The automatic end-labelling of one fragment during ribozyme cleavage is a practical advantage

Ribozyme development has only just begun The cleavage efficiency needs improving, so that digestion is complete Denaturants such as urea and formamide appear to increase the specificity of cleavage, they should also melt the secondary structure of the substrate and maximize the availability of target sequences There are 256 possible tetranucleotide cleavage enzymes of which we have investigated only three Further mutagenesis experiments are needed to discover how many of these are both efficient and specific

Protein ribonucleases cleave RNA substrates with high specificity by recognizing a combination of RNA structure and sequence, the structure is more important (for instance RNase III (ref 28) and RNase M5 (ref 29) But proteins that cleave single-stranded RNA substrates are only specific for mononucleotides or dinucleotides (for instance, RNase T_1 cleaves after guanosines[30]) Thus, the L-19 IVS RNA is considerably more sequence-specific than any known protein ribonuclease

RNA sequence recognition usually involves recognition of RNA by RNA Triplet codons in messenger RNA are recognized by anticodons in transfer RNA, and ribosome binding sites in prokaryotic mRNA are recognized using the Shine-Dalgarno interaction with 16S rRNA It is easy to imagine uses for sequence-specific protein ribonucleases in pre-mRNA splicing and 3'-end generation, but there is no evidence for such enzymes Instead, 5' splice sites in pre-mRNA are recognized using the U1 small nuclear RNA[13,31] and the 3'-end cleavage site of histone H3 pre-mRNA by U7 snRNA[32], much of the sequence discrimination is provided by complementary base-pairing Perhaps it is difficult to construct an active site in a polypeptide to bind an unstructured stretch of nucleotides specifically It is possible that in certain reactions involving RNA substrates, RNA catalysis has not been superseded by protein catalysis for the simple reason that RNA does the job better

We thank T Inoue for initiating studies of ribozyme-catalyzed GTP addition in our laboratory, O Uhlenbeck for sharing his technology for synthesis of oligoribonucleotides, B Chabot and J Steitz for providing pAK105, H Sasmor and M Caruthers for DNA synthesis, and A Sirimarco and M Neary for preparation of the manuscript and illustrations This work was supported by grant GM28039 from the NIH, grant NP-374 from the American Cancer Society, and by the National Foundation for Cancer Research

1 Kruger, K *et al. Cell* **31**, 147–157 (1982)
2 Zaug, A J, Grabowski, P J & Cech, T R *Nature* **301**, 578–583 (1983)
3 Zaug, A J, Kent, J R. & Cech, T R. *Science* **224**, 574–578 (1984)
4 Zaug, A J & Cech, T R. *Science* **231**, 470-475 (1986)
5 Zaug, A J & Cech, T R. *Biochemistry* **25**, 4478–4482 (1986)
6 Nathans, D & Smith, H O *A Rev Biochem* **44**, 273–293 (1975)
7 Davies, R. W, Waring, R. B, Ray, J A, Brown, T A. & Scazzocchio, C *Nature* **300**, 719–724 (1982)
8 Waring, R. B, Towner, P, Minter, S J & Davies, R W *Nature* **321**, 133–139 (1986)
9 Inoue, T, Sullivan, F X & Cech, T R *Cell* **43**, 431–437 (1985)
10 Garriga, G, Lambowitz, A M, Inoue, T & Cech, T R. *Nature* **322**, 86–89 (1986)
11 Been, M D & Cech, T R *Cell* **47**, 207–216 (1986)
12 Winter, G & Brownlee, G G *Nucleic Acids Res* **5**, 3129–3139 (1978)
13 Mount, S M, Pettersson, I, Hinterberger, M, Karmas, A & Steitz, J A *Cell* **33**, 509–518 (1983)
14 Zaug, A J & Cech, T R. *Science* **229**, 1060–1064 (1985)
15 Price, J V, Engberg, J & Cech, T R submitted
16 Reed, R E, Baer, M F, Guerrier-Takada, C, Donis-Keller, H & Altman, S *Cell* **30**, 627–636 (1982)
17 Guerrier-Takada, C & Altman, S *Biochemistry* **23**, 6327–6334 (1984)
18 Pace N R & Marsh, T L *Orig Life* **16**, 97–116 (1985)
19 Lowary, P, Sampson, J, Milligan, J, Groebe, D & Uhlenbeck, O C *NATO ASI Series* **110** (in the press)
20 Perea, J & Jacq, C *EMBO J* **4**, 3281-3288 (1985)
21 Bass, B L & Cech, T R. *Nature* **308**, 820-826 (1984)
22 Bass, B L & Cech, T R *Biochemistry* **25**, 4473–4477 (1986)
23 Marsh, T L & Pace N R *Science* **229**, 79–81 (1985)
24 Guerrier-Takada, C, Haydock, K, Allen, L. & Altman, S *Biochemistry* **25**, 1509-1515 (1986)
25 Szostak, J W *Nature* **322**, 83–86 (1986)
26 Waring, R B, Scazzocchio, C, Brown, T A & Davies, R W *J molec. Biol* **167**, 595–605 (1983)
27 Michel, F & Dujon, B *EMBO J* **2**, 33–38 (1983)
28 Robertson, H D *Cell* **30**, 669–672 (1982)
29 Stahl, D A, Meyhack, B & Pace, N R *Proc. natn. Acad Sci U S A.* **77**, 5644–5648 (1980)
30 Egami, F, Oshima, T & Uchida, T *Mol. Biol. Biochem. Biophys* **32**, 250–277 (1980)
31 Zhuang, Y & Weiner, A M *Cell* **46**, 827–835 (1986)
32 Birnstiel, M L, Busslinger, M & Strub, K *Cell* **41**, 349–359 (1985)
33 Donis-Keller, H *Nucleic Acids Res* **8**, 3133–3142 (1980)
34 Davanloo, P, Rosenberg, A H, Dunn, J J & Studier, F W *Proc. natn Acad Sci U S A* **81**, 2035–2039 (1984)
35 Inoue T Sullivan, F X & Cech, T R *J molec. Biol* **189**, 143–165 (1986)

LETTERS TO NATURE

Is CO_2 responsible for the outbursts of comet Halley?

P. D. Feldman*, M. F. A'Hearn†, M. C. Festou‡, L. A. McFadden‡, H. A. Weaver* & T. N. Woods*

* Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA
† Astronomy Program, University of Maryland, College Park, Maryland 20742, USA
‡ Observatoire de Besançon, 25000 Besançon, France

Outbursts are thought to originate from the nucleus of comet Halley; cameras on the Vega and Giotto spacecraft have recorded the images of jets[1,2], and rapid fluctuations are seen in the visible light curve, detected, for example, by the International Ultraviolet Explorer satellite during the time the spacecraft were making their observations[3]. This activity does not appear to be limited to when the comet is near perihelion, but has been reported for heliocentric distances >4.6 AU pre-perihelion[4]. The outbursts are probably triggered by the heating of relatively volatile material within the ice. We report here spectroscopic data from the IUE showing a strong correlation between a visible outburst in the near-nucleus region on 18.8 March 1986 and the sudden appearance of a large column abundance of CO_2^+ ions in the tail. These data suggest that CO_2 may play a significant role in the outburst process This outburst may have been the source of the detached clumps of CO^+ emission recorded by ground-based observers on 19 and 20 March 1986[5,6].

Data were obtained during a combined 16 h US1-US2 IUE shift on 18 March 1986 beginning at 11 00 UT Observations consisted of a large number of relatively short exposures taken at various offset distances from the nucleus with the long-wavelength (LWP) spectrograph to map the spatial distribution of the OH emission farther from the comet than previously[7] and thereby determine the OH scalelength An initial set of exposures was taken with all the offsets (the maximum was 8 arc min) on the sunward side of the nucleus along the Sun-comet line To examine the symmetry of the OH coma, subsequent exposures consisted of sunward-tailward pairs at equal distances, again along the Sun-comet line Three short exposures at the nucleus were also obtained, at the beginning, middle and end of the observational period Near the middle of the period, a moderately long exposure of the LWP spectrograph was carried out

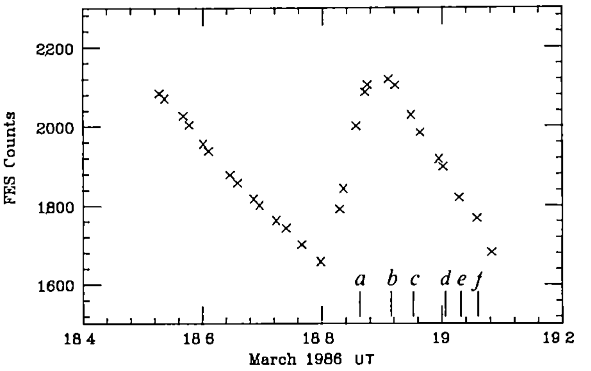


Fig. 1 Light curve of comet Halley for 18, 19 March 1986 derived from the IUE Fine Error Sensor. The effective aperture is ~18 arc s square. *a–f*, Times of the spectrograph exposures listed in Table 1.

to determine the flux of the continuum and weaker features such as CS and C_2.

The IUE fine error sensor (FES) is a slit-jaw camera used to acquire and track a desired target. The FES can also be used as a visible light photometer (wavelength range ~4,000–6,500 Å) sensitive to dust and C_2 and CN emission in a region ~18 arc s square (11,000 km × 11,000 km) around the centre of brightness of the comet. It can provide continuous monitoring of short-term variability[8] not possible at ground-based observatories. The FES data from March 18, 19 are shown in Fig 1, and indicate an outburst beginning at ~19.22 UT, increasing linearly and reaching a maximum about 2.5 h later, and then decaying at a faster rate than before the outburst. The mid-point exposure times of the spectra presented in Table 1 are also shown (*a–f*). At the time of the outburst, the heliocentric distance of Halley was 0.98 AU, the geocentric distance was 0.84 AU, the solar elongation was 63.7° and the phase angle was 66°. Foreshortening along the Sun-comet line was <10%.

Spectra *a–c* are shown in Figs 2 and 3. Spectrum *b* clearly shows the $\tilde{B}$-$\tilde{X}$ doublet of CO_2^+ at 2,884–2,896 Å, at a flux comparable to the nearby OH (1,0) band at a position 4 arc min, or 150,000 km projected on the sky, on the tailward side of the nucleus. The CO_2^+ feature is absent at the same distance on the sunward side (*c*). It is still present (despite the higher camera noise level from the shorter exposure) 1 arc min tailward in *d*, taken 2.25 h after *b*. In the spectrum at the nucleus (*a*) the CO_2^+ emission is masked by the very strong continuum of dust-scattered sunlight, and accounts for only ~15% of the observed emission at 2,890 Å. Festou *et al*[3] have suggested, from earlier IUE spectra of Halley, that this comet appears to show a much lower emission rate of CO_2^+ relative to OH than several other comets observed by the IUE[9]. However, as Fig 2*a* shows, the region of the CO_2^+ doublet contains a maximum in the spectrum of scattered solar radiation (dotted line) which masks the CO_2^+ feature. Subtraction of the solar spectrum, properly reddened[10], clearly shows the CO_2^+ doublet (*b*).

The coincidence of the tailward exposures with the outburst was extremely fortuitous, as was another observation that showed the transient nature of the outburst. The spectra denoted by *d* and *e* in Table 1 were both taken at the same offset of 1 arc min into the tail because of an operational error. A comparison shows more than a fourfold decrease in a period of 37 min for the CO_2^+ emission, whereas the OH brightness remained nearly constant. Note also the decrease in CO_2^+ brightness between spectra *b* and *d* which is all the more striking as *d* was obtained with a fourfold smaller offset into the tail.

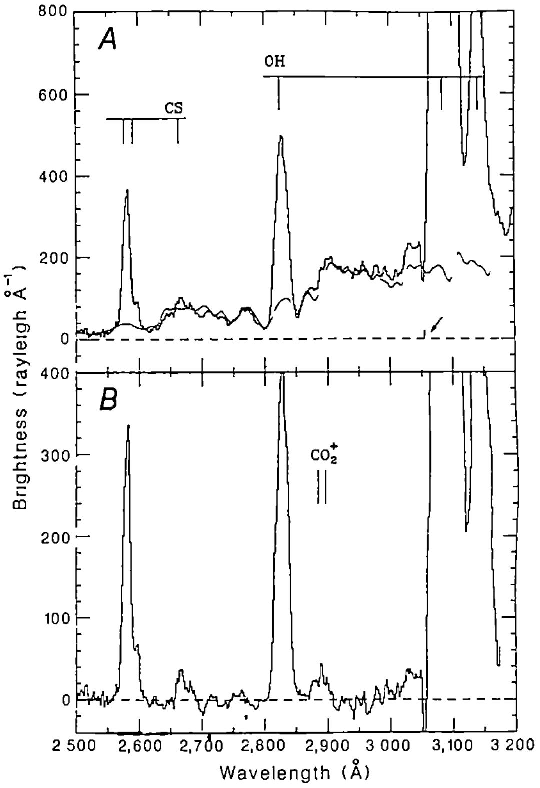


Fig 2 *A*, spectrum *a*, at the centre of brightness, showing the prominent OH and CS emission features. The dotted line is a solar spectrum smoothed to the resolution of the IUE spectrograph. The position of a camera reseau mark is indicated by the arrow. *B*, same spectrum but with a reddened solar spectrum subtracted.

Although the spatial and temporal coverage afforded by the IUE spectrograph is extremely limited, the data strongly imply a sudden enhancement followed by a rapid decrease in the abundance of CO_2^+ ions in the plasma tail in the ~10^4 s after the onset of the outburst detected by the FES. The presence of CO_2^+ suggests the production of a large enhancement of CO_2 molecules at the nucleus, perhaps enough to drive the outburst in terms of propelling grains into the coma. A qualitative estimate of the CO_2 production rate required can be derived from the observed surface brightness of the CO_2^+ emission.

The IUE spectrograph slit is ~10 arc s × 20 arc s with the long axis ~73° to the Sun-comet line for these observations. This long axis is then part of a surface 150,000 km from the nucleus through which all the ions produced inside must flow under the accelerating influence of the solar wind. The 20 arc s dimension corresponds to a linear width $w = 12{,}000$ km, and the uniform spatial distribution of the CO_2^+ emission in the aperture indicates that the width of the ion tail is considerably greater than this slit width. The average column density in the aperture, *N*, is given in terms of the measured surface brightness, *B*, in rayleighs, as

$$N = \frac{Br^2}{g} \times 10^6 \text{ cm}^{-2}$$

where *g* is the fluorescence scattering efficiency in photons per second per molecule at 1 AU and *r* is the heliocentric distance of the comet in AU.

The rate of flow of CO_2^+ ions through that portion of the surface defined by the slit is

$$q = Nvw \text{ ions s}^{-1}$$

where *v* is the outflow velocity at the surface. To estimate a lower limit to *v*, if the ions observed in spectrum *b* were produced at the nucleus at the onset of the outburst, 9,300 s earlier, their mean velocity into the tail would be 16 km s^{-1}. The assumption of uniform acceleration then gives $v > 32$ km s^{-1}. Thus, in the IUE aperture, using $g(CO_2^+) = 2 \times 10^{-3}$ photons s^{-1} ion^{-1} (at 1 AU) (ref 11 and J. L. Fox, personal communication), the observed brightness of 560 R gives $q > 1.0 \times 10^{27}$ s^{-1}. However, the width of the ion tail region is significantly greater than the 12,000 km defined by the IUE aperture, so *q* must be even larger

Table 1 IUE data from 18–19 March 1986

Sequence	Image no	Exposure duration(s)	Mid-point time of exposure (UT)	Position	$B(CO_2^+)$ (R)	B(OH 1,0) (R)	B(OH 0,0) (kR)
—	LWP7823*	30	19 56	nucleus	—	7,000	385
a	LWP7824	900	20 43	nucleus	750±260†	saturated	saturated
b	LWP7825	360	21 57	4 arc min tailward	560±20‡	645	26.2
c	LWP7826	360	22 51	4 arc min sunward	−10±25	760	31.7
d	LWP7827	90	00 08	1 arc min tailward	295±95	2,340	106
e	LWP7828	90	00 45	1 arc min tailward	60±75	2,370	105
f	LWP7829	90	01 27	1 arc min sunward	−115±80	2,060	114
—	LWP7830	30	02 03	nucleus	—	7,300	359

* This spectrum is included because the OH bands are saturated in LWP7824

† With continuum subtracted

‡ Uncertainty is standard deviation in camera noise only

Ground-based monochromatic images of other comets at similar heliocentric and geocentric distances[12], suggest that $w>$ 50,000 km, so that the ion flow rate, which is equal to the CO_2^+ production rate, must be $>4.2\times10^{27}\,s^{-1}$.

For comparison, the water production rate, derived from the observed OH (0, 0) band brightness[3], is $3.9\times10^{29}\,s^{-1}$. An upper limit to the CO_2 volume mixing ratio relative to H_2O, derived from the Giotto neutral mass spectrometer data[13], is 0.035, and if this represents a 'quiescent' value, it would imply a production rate, $Q(CO_2)<1.4\times10^{28}\,s^{-1}$. The observed CO_2^+ emission rate would then require efficient ionization and dissociation of all neutral CO_2 produced within a region $<10^4$ km from the nucleus. As the time scales for CO_2 destruction are significantly longer than 10^4 s, the only way to account for the required CO_2^+ production rate is by a large increase in $Q(CO_2)$ during the outburst over the quiescent value. Surprisingly, the OH (0,0) band emission rate does not show an increase after the outburst (see the first and last lines of Table 1), but, in fact, shows a slight decrease, despite the lifetime of the OH parent, H_2O (~1 day). Almost all the IUE data, except for this outburst, show a strong correlation between the FES count rate and the measured OH brightness[14].

The ionization source near the nucleus remains an open question. Photoionization of CO_2 by solar extreme ultraviolet photons is relatively slow (the rate at 1 AU is $\approx7\times10^{-7}\,s^{-1}$)[15] so that 10^4 s after the beginning of the outburst $<10^{-2}$ of the CO_2 molecules produced would have been ionized. Electron impact ionization is also a possible source, particularly within $\sim10^4$ km of the nucleus where significant concentrations of $10–10^3$ eV electrons were measured by the plasma experiment on Vega 1[16]. Evidence for electron impact excitation of forbidden ultraviolet transitions in the inner coma was also obtained in other ultraviolet observations of comet Halley[17]. However, although collisional ionization, in addition to photoionization, could significantly increase the CO_2^+ formation rate, at the required rate the entire neutral molecular coma (principally H_2O) would quickly be converted to ions and atoms contrary to all known observations.

Several ground-based observers[5,6] have reported monochromatic images in CO^+ ($\lambda=4,260$ Å) taken on 20 March 1986 showing a detached condensation of CO^+ at $\sim2.6\times10^6$ km from the head of comet Halley. Pederson and Vio[5] also determined the velocity of the condensation relative to fixed stars, the velocity relative to the comet nucleus was found to be $\sim35\,km\,s^{-1}$. Extrapolating backwards from their observation time of UT 20.295 March 1986, and allowing for changes in velocity of the ion cloud, it is very likely that the origin of this cloud is the outburst detected by the IUE. Unfortunately, although CO^+ is detectable in the wavelength region of the IUE exposures listed in Table 1 through fluorescence in the First Negative system near 2,200 Å (ref 18), the combination of low fluorescence efficiencies[19] and low sensitivity of the IUE camera at 2,200 Å resulted in no detection of CO^+ emission in any of the IUE spectra. However, as CO appears to be the second most abundant parent molecule in the coma of comet Halley[17] and is even more volatile than CO_2, it is likely that both species are released simultaneously during the outbursts (together with the grains that give rise to the increase in FES counts seen in Fig 1) and both the IUE and ground-based observations can be associated with a common event. The presence of CO and CO_2 is also expected on the basis of laboratory experiments that have shown that both CO and CO_2 are likely to be present in comparable abundance in pre-solar cometesimal grains[20].

The role, if any, of the solar wind in the observed outburst is not clear. Measurements of solar wind parameters by the Sakigake spacecraft[21], not too far from Halley, showed a large increase in electron density early on 18 March 1986. However, even at the higher particle flux, the rate for charge exchange ionization of molecules by solar wind protons is at best comparable to the photoionization rate, and the solar wind is stood off by mass loading from the inner coma region where most of the ionization must occur[22]. There is also no evidence in any of the sunward offset exposures taken by IUE before the outburst of

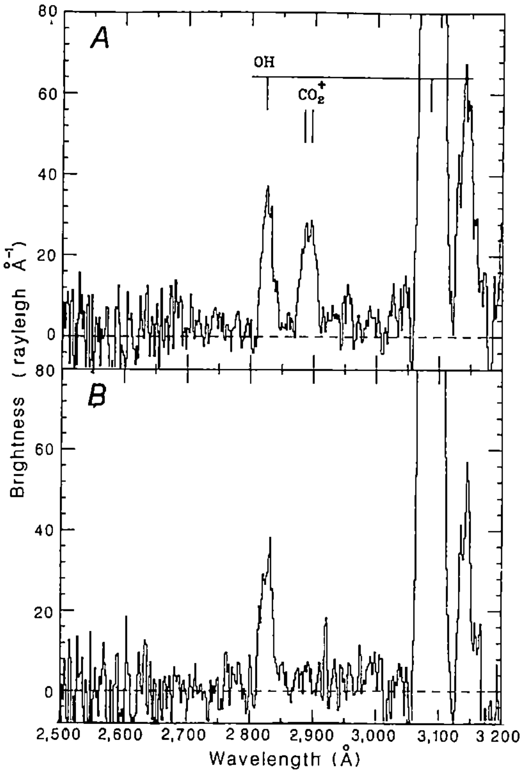


Fig 3 Spectra *b* and *c* (from Table 1) (*A* and *B* respectively) taken 4 arc min tailward and sunward, respectively

any CO_2^+ emission that would indicate enhanced ionization in the coma Thus, while the solar wind may have significantly affected the structure of the ion tail once formed, it is unlikely to have initiated the outburst

We thank the IUE staff at Goddard Space Flight Center for their assistance in the operations and acquisition of data E Roettger assisted in the data analysis This work was supported by NASA grants NSG-5393 to the Johns Hopkins University and NAG 5-252 to the University of Maryland P D F, M F A'H, M C F and H A W were guests observers with the IUE satellite observatory

Received 13 June, accepted 8 October 1986

1 Sagdeev, R Z *et al* *Nature* **321**, 262-266 (1986)
2 Keller H U *et al* *Nature* **321**, 320-326 (1986)
3 Festou, M C *et al.* *Nature* **321**, 361-363 (1986)
4 Festou, M C *et al.* *Astr Astrophys* (in the press)
5 Pedersen, H & Vio, R. *IAU Circ.* No 4198 (1986)
6 Rees, D, Meredith, N & Wallis, M *Adv Space Res* **5**, 255-262 (1986)
7 Weaver, H A, Feldman, P D, Festou, M C & A Hearn, M F *Astrophys J* **251**, 809-819 (1981)
8 Feldman, P D, A Hearn, M F & Millis, R L *Astrophys J* **282**, 799-802 (1984)
9 Festou, M C, Feldman, P D & Weaver, H A *Astrophys J* **256**, 331-338 (1982)
10. Feldman, P D & A'Hearn, M F in *Ices in the Solar System* (ed Klinger, J *et al.*) 453-461 (Reidel, Dordrecht 1985)
11 Fox, J L & Dalgarno, A *J Geophys Res* **84**, 7315-7333 (1979)
12 Rahe, J, McCracken, C W & Donn, B D *Astr Astrophys Suppl* **23**, 13-35 (1976)
13 Krankowsky, D *et al* *Nature* **321**, 326-329 (1986)
14 Feldman, P D *et al* in *New Insights in Astrophysics, 8 Years of UV Astronomy with IUE*, 39-42 (European Space Agency, SP-263, 1986
15 Huebner, W F & Carpenter, C W *Solar Photo Rate Coefficients*, LA-8085-MS, Los Alamos National Laboratory (1979)
16 Gringauz K I *et al* *Nature* **321**, 282-285 (1986)
17 Woods, T N, Feldman, P D, Dymond, K & Sahnow, D J *Nature* **324**, 436-438 (1986)
18 Feldman, P D & Brune, W H *Astrophys J* **209**, L45-L48 (1976)
19 Magnani, L & A Hearn, M F *Astrophys J* **302**, 477-487 (1986)
20 Greenberg, J M in *Cometary Exploration* Vol 2 (ed T I Gombosi) 23-54 (Hungarian Academy of Sciences, Budapest, 1983)
21 Oyama, K, Hirao, K, Hirano, T, Yumoto, K & Saito, T *Nature* **321**, 310-313 (1986)
22 Wallis, M K & Ip, W-H *Nature* **298**, 229-234 (1982)

Rocket ultraviolet spectroscopy of comet Halley and abundance of carbon monoxide and carbon

T. N. Woods, P. D. Feldman, K. F. Dymond & D. J. Sahnow

Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA

The ultraviolet spectrum of a comet's coma is dominated by emissions from the dissociation products of water, OH, H and O, and from secondary species: C, C^+, CO, CO^+, CO_2^+, S, and CS (ref. 1) We report here two far ultraviolet observations of comet Halley made on 26 February 1986 and 13 March 1986 with a sounding rocket experiment. This is the first time that long-slit spectroscopy, a standard technique for the study of extended objects such as comets and nebulae, has been applied to the far ultraviolet study of a comet. The observed CO spatial profiles can be modelled by a radial outflow model[2] for a parent molecule and suggest that the CO is produced directly from the nucleus of the comet Using the observed O I emission profile to deduce the H_2O production rate, the abundance of CO relative to H_2O is found to be 20±5% for the first flight and 17±4% for the second flight, making CO the second most abundant parent molecule in the coma The derived production rate of atomic carbon is consistent with that expected from the photodissociation of carbon monoxide.

Atomic carbon is a common feature in the ultraviolet spectrum of many comets, yet its origin remains unclear In comet West (1976 VI), the observed C I emission was consistent with a model of carbon as a daugher of CO which is vaporized from the nucleus as a parent molecule[3] However, for comet Bradfield (1979X) the C I brightness and spatial distribution were not consistent with models of carbon either as the daughter of the observed CO or as the granddaughter of CO_2 (ref 4) The differences may reflect intrinsic compositional differences between comets, or deficiencies in the models used to interpret the observations Festou[5] has discussed this 'carbon puzzle' in the context of similar results on other comets observed by the International Ultraviolet Explorer (IUE) satellite

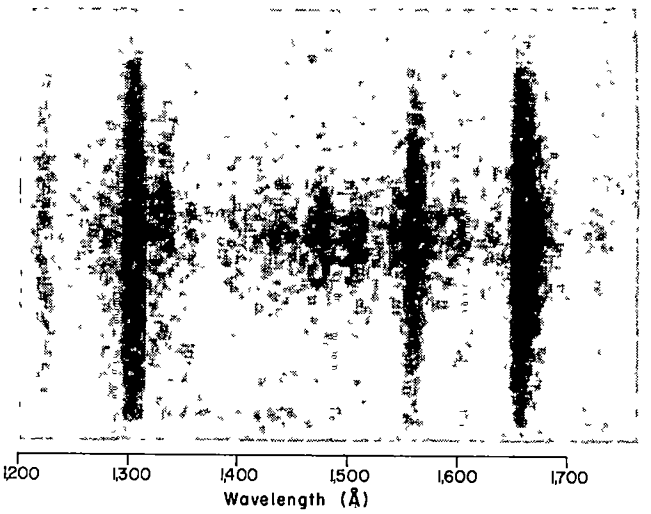


Fig 1 A photographic representation of the raw counts from the spectrograph on NASA 21 093UG is shown for a 72 2-s exposure of comet Halley centred in the entrance slit of the spectrograph The horizontal axis is the wavelength dispersion from ~1,150 to 1,800 Å, and the vertical axis is along the 7 7 arc min slit with the sunward direction pointing down The three brightest lines are O I λ1,304, C I λ1,561, and C I λ1,657 H I λ1,216 appears as a weak feature due to the heavy attenuation below 1,230 Å by a CaF_2 filter in front of the entrance slit

The spectral range for the sounding rocket observations, 1,200-1,750 Å at 12 5 Å resolution, is well suited for studying the carbon chemistry of the coma as it includes C I λ1,657, C I λ1,561, C II λ1,335, and the CO fourth positive system whose strongest bands are at 1,478 and 1,510 Å. While these observations augment the many ultraviolet observations of comet Halley made with IUE[6], the sounding rocket payload is specially designed for the study of the emission profiles in the coma because the spectrograph provides 10 arc s spatial resolution along the 7 7 arc min slit In addition, a higher sensitivity, needed because of the short observing time (~300 s) available with a sounding rocket experiment, is achieved using a photon-counting detector

The faint object telescope (FOT) payload flown previously[7] was modified for long-slit spectroscopy of extended emission objects The basic structure of the FOT is a Dall-Kirkham telescope, a Rowland circle spectrograph and a slit-jaw camera The spectrograph modifications included the use of an ellipsoidal grating to reduce astigmatism at the focal plane and a two-dimensional resistive anode detector The first launch of the experiment, NASA 21 093UG, was on 26 February 1986 at 12 02 UT, 17 days after perihelion, when the heliocentric distance, *r*, was 0 70 AU and the geocentric distance, Δ, was 1 32 AU The payload was recovered, refurbished, and flown again on 13 March 1986 at 11 20 UT when *r* was 0 90 AU and Δ was 0 98 AU This flight, NASA 21 095UG, preceded the Giotto encounter by ~13 h Both launches were from the White Sands Missile Range, New Mexico (106 3° W, 32 4° N)

During both flights a central exposure was taken with the entrance slit (16 arc s wide by 7 7 arc min long) of the spectrograph centred on the brightest part of the coma and with the long axis of the slit oriented along the Sun-comet line The first flight also included a tail exposure that was obtained with the slit centred 8 arc min away from the comet in the anti-sunward direction, again with the slit along the Sun-comet line

Table 1 The brightness *B* and the derived column density *N* of the species in the central coma (10×16 arc s) observed in both flights

	NASA 21 093 UG 26 February 1986, 12 07 UT			NASA 21 095 UG 13 March 1986, 11 25 UT			IUE
Species	*B** (R)	*g*† (1 AU) (10^{-6} s^{-1})	*N* (10^{14} cm^{-2})	*B** (R)	*g*† (1 AU) (10^{-6} s^{-1})	*N* (10^{14} cm^{-2})	*B*‡ (R)
O I λ1,304	1,460±320	3 7	1 9	280±65	1 8	1 25	210±40
C II λ1,335	130±30	50 0	0 013	25±15	37 0	0 0054	—
CO λ1,478	175±50	0 24	3 5	65±25	0 24	2 2	100±60
CO λ1,510	140±40	0 18	3 8	60±25	0 18	2 7	60±40
C I λ1,561	480±110	6 0	0 39	150±35	5 9	0 20	190±50
C I λ1,657	1,600±350	23 0	0 34	520±115	22 0	0 19	745±110

* The uncertainty is a combination of the 20% uncertainty in the absolute calibration of the instrument and the counting statistics

† The solar fluorescence efficiencies (or *g*-factor) are based on earlier values[14] which are adjusted for the heliocentric velocity of comet Halley and the solar flux (G Rottman, personal communication) during the observation The CO *g*-factors are from ref 8 scaled to 1 AU and corrected by 0 5 to account for the variation of the solar flux

‡ The average of IUE data from 12 and 13 March 1986, the CO brightness is from 13 March only

A spectrogram of the central exposure taken during the first flight is represented as a two-dimensional photographic image in Fig 1 The vertical axis depicts the spatial distribution of the emissions from the coma The extended and brighter features in the spectrogram are H I λ1,216, O I λ1,304, C I λ1,561, and C I λ1,657, which have been observed in almost all comets studied in the far ultraviolet region[1] The spectra of the central coma (16×72 arc s centred on the nucleus) are shown in Fig 2 for both flights Other features identified in the spectra include the emissions from C II λ1,335 and the CO fourth positive system At least 11 bands of CO are clearly identified by comparison with a synthetic spectrum based on CO fluorescence of ultraviolet solar radiation[8] While a few of these bands are seen in the IUE spectra of comet Halley[6], the poor signal-to-noise ratio in the IUE spectra makes it difficult to extract reliable band intensity ratios from those data

A tentative identification of O I λ1,356 at a level of 60 R in the 26 February spectrum can be made This intercombination line is not excited by solar fluorescence as the other emissions are, and suggests the presence of a region of electron excitation within the inner coma Other indications for a collisional destruction mechanism within ~10^4 km of the nucleus besides solar photodissociation and photoionization are given by the IUE observations of CO_2^+ during an outburst[9] and by the unexpectedly high C^+ abundance measured by Giotto[10] An enhanced flux of keV electrons at 15,000 km measured by Vega 1 and Vega 2[11] and a maximum in the total ion current at 12,000 km as measured by Giotto[12] indicate that a region in the inner coma could exist where collisional excitation or destruction mechanisms are significant Additional modelling is needed to verify this identification of O I λ1,356

The brightnesses of the strongest emission features for both flights are listed in Table 1, and are given for an area of 10×16 arc s centred on the nucleus of comet Halley Table 1 also gives the derived column densities and the fluorescence efficiencies used to derive them The IUE data in Table 1, given for comparison with NASA 21 095UG, are the average of observations made on 12 and 13 March 1986 at 22 UT on each day The IUE data are reduced to an effective aperture of 10×15 arc s These brightnesses are consistent within the uncertainty in the two measurements, which results from a combination of calibration errors and the different spatial resolution and tracking stability of the two instruments Although the visual brightness of the comet varied by a factor of two over a period of a day during many of the IUE observations[6], the O I and C I emissions should not vary as rapidly as their parents' lifetimes are longer than a day Because the visual brightness was reasonably constant from 12 to 13 March, only a small effect due to the intrinsic variability is expected in comparing the two data sets

As the composition of the comet's nucleus cannot be directly inferred from the given brightnesses, modelling of the coma to fit the observed radial distributions of the emissions is performed to determine the production rates of the molecular species A radial outflow model of CO with a production rate of 2 4×10^{29} s^{-1} for the first flight and 1×10^{29} s^{-1} for the second flight yields a good fit to the observed CO distributions (Fig 3*a*) This

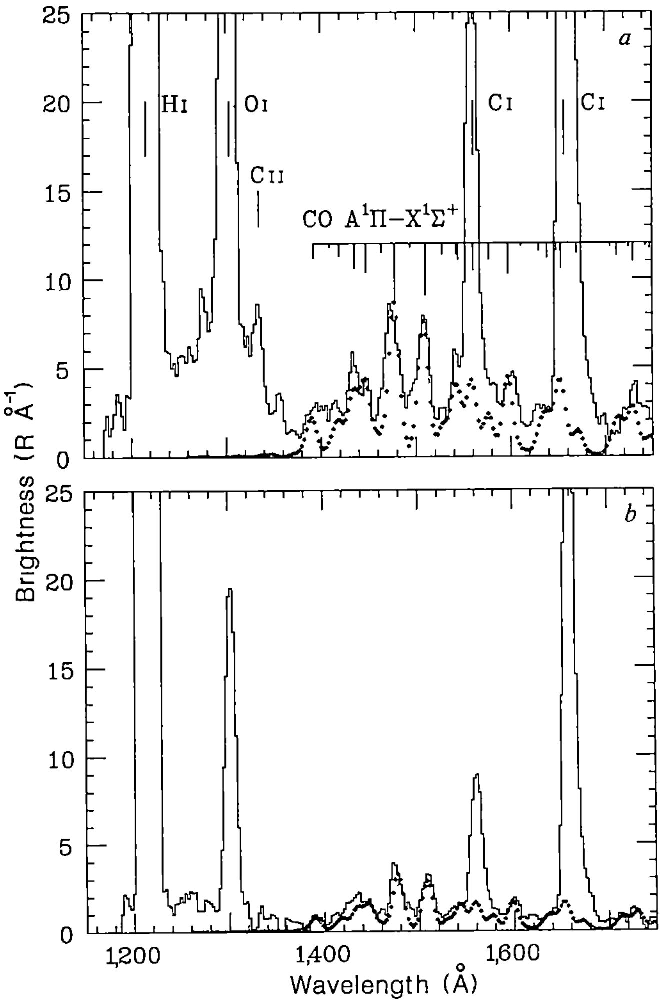


Fig 2 The central coma spectrum from *a*, the first flight on 26 February 1986, and *b*, the second flight on 13 March 1986 The synthetic CO spectrum (diamonds), convolved to our instrument resolution, is based on the *g*-factors of CO fluorescence of ultraviolet solar radiation given by Durrance[8] The feature at 1,356 Å in *a* has been tentatively identified as O I λ1,356 excited by electron collisions in the inner coma

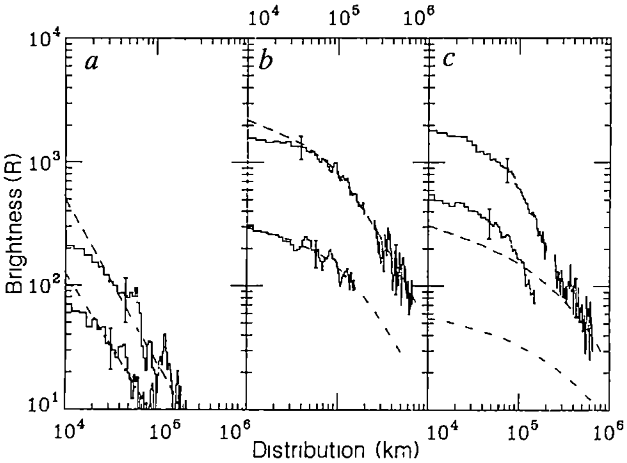


Fig. 3 The observed radial distribution of *a*, the 1,478-Å CO band, *b*, O I λ1,304, and *c*, C I λ1,657 for both flights with the data from the first flight being the upper curve in each panel. The dashed lines are emission profiles, for comparison with the data from the first flight, convolved to our instrument resolution, from a radial outflow model with the production rate of $2.4\times10^{29}\,s^{-1}$ for CO and $1.2\times10^{30}\,s^{-1}$ for H_2O. The dot-dash lines, for comparison with the data from the second flight, are emission profiles from a model with the production rate of $1\times10^{29}\,s^{-1}$ for CO and $6\times10^{29}\,s^{-1}$ for H_2O.

strongly suggests that the CO is vaporized directly from the nucleus, but cannot exclude the possibility of a parent species which has a lifetime <1,000 s. A possible parent for CO is H_2CO, but it is improbable because a radial outflow model of H_2CO that reproduced the observed CO profile would require a production rate greater than the water production rate and would produce a detectable level of H_2 fluorescence in our spectral range. By fitting an outflow model with H_2O and CO parents to the observed oxygen distributions in Fig. 3*b*, a production rate for water of $1.2\times10^{30}\,s^{-1}$ is derived for the first flight and $6\times10^{29}\,s^{-1}$ for the second flight. The average water production rate derived from the OH (0, 0) band brightness measured by IUE, using a vectorial model[2], is $5\times10^{29}\,s^{-1}$ for comet Halley on 12 and 13 March 1986. The water production rate derived from the Giotto neutral mass spectrometer data[12] is $5.5\times10^{29}\,s^{-1}$ with an estimated 50% uncertainty. Because of the uncertainties in the parameters (lifetime and velocity of each species) used in the model, the H_2O production rates derived from a coma model have at least a 30% uncertainty. Therefore, these water production rates are consistent within the uncertainties in the models, instrumental calibrations and the solar fluxes used in deriving the fluorescence efficiency for O I λ1,304 (G. Rottman, personal communication). Note that because CO appears to be a parent molecule and is modelled with fewer parameters, the uncertainty in the derived CO production rate is smaller than the uncertainty in the H_2O production rate.

While data from the Giotto ion mass spectrometer gave an upper limit of 20% for the ratio of CO to H_2O abundance[10], and a rough estimate from the IUE observations in March for this ratio is 10-20%[6], the results from the rocket data give a relative abundance of CO to H_2O as 20±5% for the first flight and 17±4% for the second flight. The uncertainty in these ratios includes the errors from the absolute calibration of the instrument and the fit of the models to the observed radial profiles. Krankowsky *et al.*[12] consider CO_2, NH_3 and CH_4 to be the second most abundant molecules in the coma, however, our results indicate that CO is the second most abundant parent molecule with a mixing ratio of ~18% relative to H_2O. For comparison, a 27% abundance of CO to H_2O was derived for comet West[3], and 1% for comet Bradfield[4].

While these data strongly suggest that the CO is vaporized directly from the nucleus, an unresolved difference remains between the observed radial distribution of carbon shown in Fig. 3*c* and the predictions of the CO outflow model. The data seem to require an additional source of carbon. An outflow model with a small additional source of carbon from the nucleus, at ~2% of the H_2O production rate, improves the fit to the observed carbon profile, but it is not adequate to identify firmly a direct source of carbon from the nucleus. An alternative evaluation of the atomic carbon production rate can be obtained if the total flux in a C I emission multiplet is known. Although the present experiment did not measure the total flux, a reliable estimate can be obtained from the observed radial profiles because the measurements in the tailward exposure on the first flight extend to 7×10^5 km. Circular symmetry of the carbon emission, necessary for this derivation, is a safe assumption as the observed carbon emissions are symmetrical on the sunward and anti-sunward sides of the nucleus. The derived carbon production rate, Q_C, is $4.3\times10^{28}\,s^{-1}$ for the first flight. As the carbon scale length is ~2.5×10^6 km at 0.70 AU, this production rate is only a lower limit. Extrapolating the observed carbon radial distribution to 3×10^6 km yields $Q_C=6.4\times10^{28}\,s^{-1}$, while extrapolation to 1.5×10^7 km gives $Q_C=6.6\times10^{28}\,s^{-1}$. Thus, a reasonable estimate for the total carbon production rate for the first flight is $7\times10^{28}\,s^{-1}$, which gives a ratio of Q_{CO}/Q_C of 3.4±1.2. Because our derived carbon production rate is a lower limit, we consider this ratio to be consistent with the expected ratio of 2.1, assuming that C is produced only through the photodissociation of CO (ref. 13).

The apparent contradiction between the need for an additional carbon source and the result that the Q_{CO}/Q_C ratio is consistent with photodissociation of CO may be resolved by the inclusion of an electron impact source in the models. Evidence for such a source, including the OI λ1,356 emission in Fig. 2, has been cited above. If electron impact enhanced the dissociation of CO in the inner coma, without destroying a significant fraction of the total CO, then the radial distribution of C would exhibit a steeper slope in agreement with the observations, while the radial distribution of CO and the Q_{CO}/Q_C ratio would not be significantly altered. Quantitative modelling using the *in situ* plasma measurements[10,11] is needed to evaluate the possible importance of this mechanism. In addition, improved coma models, such as the vectorial model[2], or models based on the *in situ* measured abundances of CO_2, CH_4, and other carbon-bearing molecules may also provide better agreement with the observed radial distributions of the carbon emissions.

We thank J. van Overeem and the staff of the NASA Sounding Rocket Division, P. Meredith and C. Jenkins of the Aerojet Techsystems Company, R. Pelton of Johns Hopkins University, and the research rocket group at White Sands Missile Range for support and field assistance. We thank G. Rottman and B. Knapp of the University of Colorado for providing unpublished solar flux data from the Solar Mesosphere Explorer satellite. This work was supported by NASA grant NAG 5-619.

Received 13 June; accepted 10 September 1986.

1. Feldman, P. D. *Science* **219**, 347-354 (1983).
2. Festou, M. C. *Astr. Astrophys.* **95**, 69-79; **96**, 52-57 (1981).
3. Feldman, P. D. *Astr. Astrophys.* **70**, 547-553 (1978); in *Asteroids, Comets Meteors II* (eds Lagerkvist, C.-I. *et al.*) 263-267 (Reprocentralen HSC, Uppsala, 1986).
4. Weaver, H. A. thesis, The Johns Hopkins Univ. (1981).
5. Festou, M. C. *Adv. Space Res.* **4**, 165-175 (1984).
6. Festou, M. C. *et al.* *Nature* **321**, 361-363 (1986).
7. Hartig, G. F., Fastie, W. G. & Davidsen, A. F. *Appl. Opt.* **19**, 729-740 (1980).
8. Durrance, S. T. *J. geophys. Res.* **86**, 9115-9124 (1981).
9. Feldman, P. D. *et al.* *Nature* (in the press).
10. Balsiger, H. *et al.* *Nature* **321**, 330-334 (1986).
11. Gringauz, K. I. *et al.* *Nature* **321**, 282-285 (1986).
12. Krankowsky, D. *et al.* *Nature* **321**, 326-329 (1986).
13. Huebner, W. F. & Carpenter, C. W. *Solar Photo Rate Coefficients* (Los Alamos Scientific Laboratory Rep. LA 8085-MS, 1979).
14. Feldman, P. D. in *Comets* (ed. Wilkening, L. L.) 461-479 (University of Arizona Press, Tucson, 1983).

Far-ultraviolet spectral images of comet Halley from sounding rockets

R. P. McCoy*, C. B. Opal† & G. R. Carruthers*

* E. O. Hulburt Center for Space Research,
Naval Research Laboratory, Washington, D.C. 20375-5000, USA
† McDonald Observatory, Department of of Astronomy,
University of Texas, Austin, Texas 78712, USA

Wide-field ultraviolet imagery of comets provides a useful means of determining such fundamental parameters as production rates, lifetimes and outflow velocities of daughter photodissociation products in the coma. Far-ultraviolet imagery has been used to measure the production rates and outflow characteristics of hydrogen for comets Bennett[1], Kohoutek[2,3] and West[4,5], for C and O in comet Kohoutek[6], and for C, S and CS in comet West[7]. Ultraviolet spectroscopy has also revealed the presence of C^+ ions and CO molecules[8] in comet West. Other comets have been observed spectroscopically in the ultraviolet from the International Ultraviolet Explorer (IUE) spacecraft[9]. Here we present far-ultraviolet images of comet Halley obtained from sounding rockets launched from White Sands Missile Range, New Mexico, on 24 February and 13 March 1986. We gained direct electrographic images of the hydrogen coma of the comet at the Lyman-α wavelength (1,216 Å), and objective spectra containing images of the coma at the oxygen (1,304 Å), carbon (1,561 and 1,657 Å) and sulphur (1,814 Å) resonance multiplets. Analysis of the Lyman-α images yields hydrogen atom production rates of $1.9\times10^{30}\ s^{-1}$ and $1.4\times10^{30}\ s^{-1}$ for the two observations. Images of oxygen, carbon and sulphur emissions obtained with the objective grating spectrograph are presented for the first set of observations and preliminary production rates are derived for these elements.

Preliminary results of two far-ultraviolet rocket imaging experiments are reported here. The first rocket (NASA 36 010 DL), was launched as soon after perihelion as practical to observe the comet at high luminosity. The payload was refurbished at the Naval Research Laboratory between launches. The second rocket (NASA 36 017 DL) was launched just before the Giotto spacecraft flyby of comet Halley. Both experiments were coordinated with rocket observations of comet Halley by Johns Hopkins University experimenters (accompanying paper)[10]. The relative solar and cometary positions at the time of the two flights are given in Table 1.

The rocket payload instruments were similar to those used previously in sounding rocket observations of comet Kohoutek[2]. The payload contained a direct-imaging camera covering the wavelength range 1,050–1,600 Å (which includes the hydrogen Lyman-α line at 1,216 Å), and an objective grating spectrograph covering the range 1,230–2,000 Å. Also included was a Lyman-α airglow calibration photometer using a photoionization chamber filled with nitric oxide as a detector.

Table 1 Observation parameters

	36 010 (24 February 1986)	36 017 (13 March 1986)
Time of launch	11 53 30 UT	10 54 00 UT
Apogee (km)	312	312
Comet zenith angle (apogee)	91°	87°
Solar zenith angle (apogee)	112°	120°
Comet-Sun distance (AU)	0.672	0.895
Comet-Earth distance (AU)	1.356	0.975
Comet-Sun velocity (km s^{-1})	17.97	25.57
Comet-Earth velocity (km s^{-1})	−32.11	−43.32
Sun-comet-Earth angle	44.0°	64.1°
Sun-Earth-comet angle	28.2°	54.1°

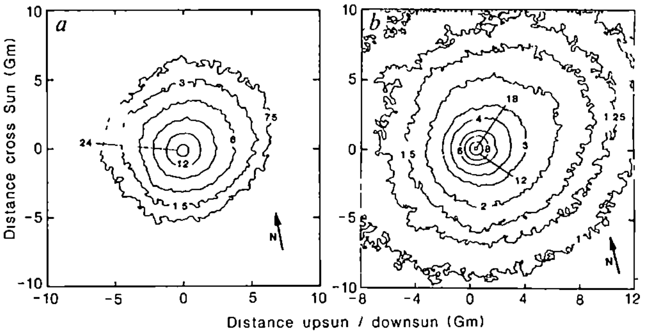


Fig. 1 Isophote contours (kR) of Lyman-α images of comet Halley. Exposure times are 2.5 s for the 24 February (a) and 9.5 s for the 13 March (b) data. The distance scales are the same and the Sun is to the left. The dotted contours on the 24 February plot are reconstructions of isophotes marred by static discharges on the film. Backgrounds due to the terrestrial Lyman-α airglow of an effective 7 kR for the first observation and 6.5 kR for the second have been subtracted.

The direct imaging camera had a 75-mm focal length, $f/1$ Schmidt optical system (LiF corrector plate) and a KBr photocathode, it was identical to Lyman-α cameras used in previous rocket observations of comets Kohoutek and West[2-4]. The camera field of view was 20° in diameter with an angular resolution of ~3 arc min.

The objective grating spectrograph was based on an electrographic Schmidt camera with a focal length of 150 mm, $f/1.5$ focal ratio, CaF_2 corrector plate and CsI photocathode. A objective grating with 1,200 lines per mm was used that gave a spectral dispersion at the photocathode of 30 Å mm^{-1}. The spectrograph had a 12° square field of view and angular resolution of ~1 arc min. The corresponding point-source spectral resolution was ~1 Å. This camera was much more sensitive than that used in the previous objective spectroscopy of comet Kohoutek because it incorporated a microchannel plate (MCP) intensification stage[11,12]. The MCP was operated at a gain of ~100, giving a corresponding increase in effective diffuse-source sensitivity.

Laboratory calibrations consisted of measurements of relative wavelength sensitivity in the 1,150–1,850 Å range combined with absolute measurements of a monochromatic 1,304 Å spot source of diameter 0.27°. For the Lyman-α camera, in-flight calibration was obtained from the ion chamber, and a relative calibration between the two flights was obtained from simulated aperture photometry of the image of the star π Capricornis, which in both cases was in the field of view. Laboratory and in-flight calibrations agreed closely, and the estimated absolute accuracy of the Lyman-α measurements is ±20%. The reduction of the laboratory calibration data for the spectrograph data is complex and has not been completed, consequently an in-flight calibration based on the spectrum of τ Capricornis, which was adjacent to the comet spectrum, was used. The observed τ Cap spectrum was calibrated using the OAO-2 photometer measurements at 1,910 Å (ref. 13), to which was normalized an OAO-2 spectrum of the star ζ Draconis[14] which has the same spectral type and colour (B6 III, $(B-V)=-0.12$). The wavelength scale in the stellar spectrum was based on an IUE spectrum of HD 182255 (ref. 15), that also has the same spectral type and reddening. The TD-1 spectrophotometry of τ Capricornis[16] is available down to 1,380 Å, but is about 15% below the OAO-2 measurement at 1,910 Å. The absolute accuracy of the spectral data is ~±35%, including the uncertainty in the comparison star spectrophotometry, but the relative (wavelength) accuracy is ~±10%.

The flight and calibration exposures were scanned and digitized using a microdensitometer at the University of Texas, Austin, using a 16×16 μm aperture on a 2048×1536 array. The measurements were then summed on a 2×2 basis to form a

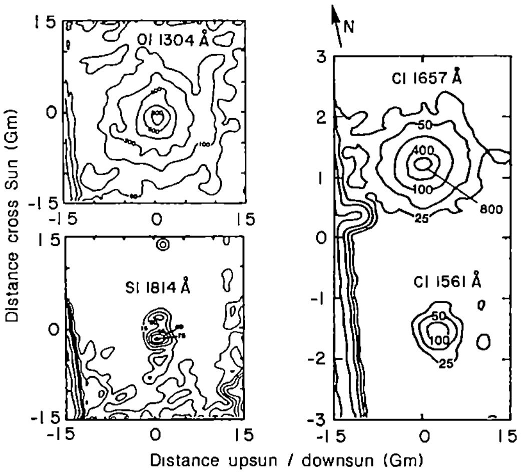


Fig. 2 Brightness in rayleighs of the objective spectral images of C, O and S resonance multiplets seen during the flight of 24 February. The images have been compressed in the spectral direction to correct for the non-isotropic plate scale of an objective spectrograph. The brightness contours have been oriented so that the Sun is positioned to the left. A background derived from an average in the vicinity of each feature was subtracted from all four images. The vertical streaks on the left are from the spectrum of τ Capricornus which was used for calibration. The right sides show a small amount of contamination from another stellar spectrum, and contours from zero order star images are seen.

1,024×768 array for final analysis, the resultant sampling was 2–3 pixels per resolution element. Nonlinearities are small in the electrographic process below a film density of ~4.0, and were avoided by using data from shorter exposures for the brighter features.

Images of the Lyman-α coma of comet Halley were obtained on both flights with 0.5, 2.5, 9.5 and 29.5 s exposure times. Contour plots of the Lyman-α isophotes from images obtained on the two periods of observation are given in Fig. 1. Only the 0.5 and 2.5 s exposures from the first flight could be used because of instrumental difficulties.

The intensity contours show approximately circular isophotes within 2 Gm of the nucleus but also show noticeable elongation away from the Sun at increasing distance. The surface brightness during the February observations was higher, but the outer isophotes of the March image are better developed due to the longer exposure time and decreased Earth-comet distance.

In both rocket flights, objective spectral images of the comet were obtained with exposure times of 2.5, 9.5 and 29.5 s. Grayscale reproductions of images from the first flight have been published[17]. Contours of the surface brightness of the O 1,304 Å, C 1,561 Å and 1,657 Å, and S 1,814 Å multiplets are shown in Fig. 2. Because the inner regions of the O 1,304 Å and C 1,657 Å images were saturated on the film for exposure times >2.5 s, the isophote contours shown for these features are a combination of images with long and short exposure times. Isophotes for contours >100 R were taken from 2.5 s exposures for both features. The outer isophotes of the 1,304 Å image were taken from a 9.5 s exposure and those of the 1,657 Å images were taken from a sum of six 29.5 s images. The C 1,561 Å and S 1,814 Å isophotes were also taken from the sum of six images.

The production rates of C, O, S and H were calculated using a simple radial-outflow model similar to that used previously to analyse hydrogen imagery[2,4]. Atoms are assumed to expand outwards isotropically from a negligibly small source region. The atoms are ultimately lost to photoionization or charge exchange with the solar wind, however, in the inner regions losses can be ignored and the column density of atoms at projected distance ρ from the nucleus can be related to the production rate Q, and mean outflow velocity $\bar{v}$, by

$$N(\rho) = Q/(4\rho\bar{v})$$

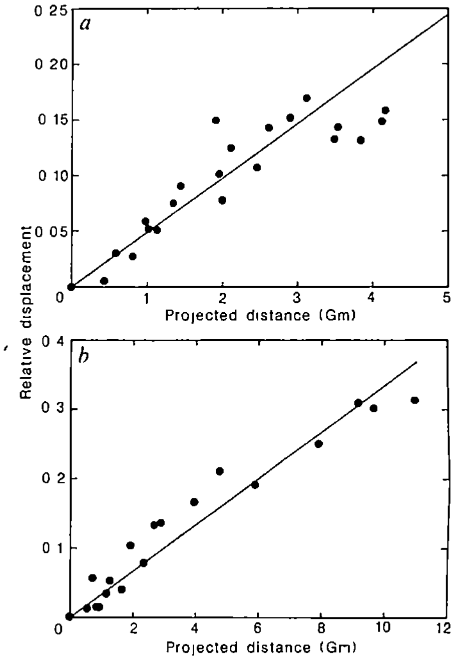


Fig. 3 Relative shift of isophote centre, S/ρ, versus ρ, the projected radius of the isophote. *a*, 24 February; *b*, 13 March. The lines show the predictions for a simple radial outflow model with $\bar{v} = 8\ \mathrm{km\ s^{-1}}$ and the acceleration due to solar Lyman-α radiation pressure produced by a solar flux of 2.3×10^{11} photons $\mathrm{cm^{-2}\ s^{-1}\ Å^{-1}}$ at 1 AU.

The brightness of the comet in rayleighs at ρ is

$$B(\rho) = 10^{-6}\, N(\rho)\, (g_o/R^2)$$

where g_0 is the g-factor, or scattering probability per (atom-second) at 1 AU, and R is the Sun-comet distance in AU. Thus, if the mean outflow velocity is known, the production rate can be determined by measuring the brightness at a distance ρ that is large compared to the instrument resolution but small compared to the scale length of the loss processes.

The radial outflow velocity for heavy constituents is typically $1\ \mathrm{km\ s^{-1}}$, and is the value adopted here for C, O and S to be consistent with previous calculations. For hydrogen, because of its small mass, a high outflow velocity is imparted during dissociation which can range from a few $\mathrm{km\ s^{-1}}$ to over $20\ \mathrm{km\ s^{-1}}$ depending on the dissociation channel. A mean outflow velocity for H can be determined by measuring the shift of the inner Lyman-α isophotes caused by solar radiation pressure[2,4]. The shell of hydrogen emitted at time t before the observation will have been accelerated downsun a distance $S = at^2/2$, where a is the acceleration due to radiation pressure. Because of limb brightening most of the contribution to the observed radiation from a given shell is from a ring of projected distance equal to the radius $\rho = \bar{v}t$. A plot of relative isophote shift versus radius then gives the velocity (Fig. 3). The prediction for $\bar{v} = 8\ \mathrm{km\ s^{-1}}$ is shown, this is in good agreement with results obtained for comets Bennett, Kohoutek and West. In spite of the many simplifying assumptions, this model tends to give production rates similar to those obtained from more realistic computations.

Production rates derived for the four species on 24 February and for H on 13 March are summarized in Table 2. The formula above was used, taking the average brightness of the coma at the radius indicated in Table 2 and assuming an outflow velocity of 1 or $8\ \mathrm{km\ s^{-1}}$ as appropriate. The sulphur brightness listed is the sum from the two bright lines that are resolved, and is the least reliable because of the poor signal-to-noise ratio for this multiplet.

The Lyman-α observations on 13 March give production rates of H_2O, (assuming all H comes from water), of $7\times10^{29}\ \mathrm{s^{-1}}$, in

Table 2 Production rates

Species	ρ (Gm)	$B(\rho)$ (R)	g_0 (1 AU) (s^{-1})	Outflow velocity (km s^{-1})	Q (s^{-1})
H (24 Feb)	2 0	7,850	$1\ 22\times10^{-3}$	8	$1\ 9\times10^{30}$
O	0 2	406	$3\ 7\times10^{-6}$	1	$5\ 9\times10^{29}$
C	0 2	535	$2\ 3\times10^{-5}$	1	$8\ 4\times10^{28}$
S	0 12	51	$5\ 9\times10^{-5}$	1	$2\ 4\times10^{27}$
H (13 Mar)	1 0	6,680	$1\ 22\times10^{-3}$	8	$1\ 4\times10^{30}$

Values for g_0 for C and O from ref 21, g_0 for H from ref 4, g_0 for S from ref 7 The 1,657 Å emission was used for the C production rate

good agreement with results from optical and mass spectrometers on the Vega[18,19] and Giotto[20] spacecraft The 24 February observations given an H/O ratio of >3, perhaps indicating that not all the hydrogen is produced from water (uncertainties could be as large as 50% for O) Our C production rate is consistent with results from the other rocket observations[10], and indicates that gaseous carbon is relatively scarce in this comet Our C/S ratio is not very different from the value of 22 given by ref 7 The cosmic ratio is 21

Further reduction of the data is in progress Future analyses of the spectral images will involved the use of more complex models, and comparisons will be made with Lyman-α observations from the Suisei, IUE, Dynamics Explorer and Pioneer Venus spacecraft

We thank D King and H Merchant for assistance with the development of the experiment payload, and the NASA Wallops payload team and the research rocket group at White Sands Missile Range for their assistance Solar Lyman-α fluxes were provided by G Rottman Supported by NASA grant S21601D

Received 9 July, accepted 30 September 1986

1 Code, A D , Houck, T E & Lillie, C F in *The Scientific Results from Orbiting Astronomical Observatory* (*OAO-2*) (ed Code, A D) 109-114 (NASA SP-310, Washington D C 1972)
2 Opal, C B , Carruthers, G R , Prinz, D K & Meier, R R *Science* **185**, 702-705 (1974)
3 Meier, R R , Opal, C B , Keller, H U , Page, T L & Carruthers, G R *Astr Astrophys* **52**, 272-281 (1976)
4 Opal, C B & Carruthers, G R *Icarus* **31**, 503-509 (1977)
5 Keller, H U & Meier, R R *Astr Astrophys* **52**, 272-281 (1980)
6 Opal, C B & Carruthers, G R *Astrophys J* **211**, 294-299 (1977)
7 Smith, A M , Stecher, T P & Casswell, L *Astrophys J* **242**, 402-410 (1980)
8 Feldman, P D & Brune, W H *Astrophys J* **209**, L145-L148 (1976)
9 F Feldman, P D in *Comets* (ed Wilkening, L L) 461-479 (University of Arizona Press, Tucson Arizona 1982)
10 Woods, T N , Feldman, P D , Dymond, K & Sahnow D *Nature* **324**, 436-438 (1986)
11 Carruthers, G R *Proc. SPIE* **172**, 304-316 (1979)
12 Heckathorn, H M & Carruthers, G R *Proc. SPIE* (*Instr in Astron. IV*) **331**, 415-426 (1982)
13 Code, A D , Hom, A V & Bottemiller, R L *Astrophys J Suppl.* **43**, 501-545 (1980)
14 Code, A D & Meade, M R *Astrophys J Suppl.* **39**, 195-289 (1979)
15 Wu, C C *et al. IUE Circ.* No 22 (1983)
16 Jamar, C *Ultraviolet Bright-star Spectrophotometric Catalogue* (*ESA SR 27* (European Space Agency, November 1976)
17 Brandt, J L *Nature* **321**, 391-393 (1986)
18 Krasnopolsky, V A *et al. Nature* **321**, 269-271 (1986)
19 Moreels, G *et al Nature* **321**, 271-273 (1986)
20 Kranskowsky, D *et al. Nature* **321**, 326-329 (1986)
21 Feldman, P D , Opal, C B , Meier, R R & Nicolas, K. R in *The Study of Comets* (eds Donn, B , Mumma, M Jackson, W , A'Hearn, M Harrington, R.) 773, 795 (NASA SP-393, Washington, D C , 1976)

Post-perihelion observations of water in comet Halley

H. A. Weaver[*][||], M. J. Mumma[†], H. P. Larson[‡] & D. S. Davis[§]

[*] The Center for Astrophysical Sciences, The Johns Hopkins University, Homewood Campus, Baltimore, Maryland 21218, USA
[†] Planetary Systems Branch, Code 693, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA
[‡] The Lunar and Planetary Laboratory, The University of Arizona, Tucson, Arizona 85721, USA
[§] Steward Observatory, The University of Arizona, Tucson, Arizona 85721, USA

Our previous pre-perihelion observations[1] of comet Halley from the NASA Kuiper Airborne Observatory (KAO) resulted in the first definite detection of water vapour in a comet. This success allowed us to carry out an ambitious post-perihelion KAO observing programme which significantly extended the scope of our Halley water investigation. The spectral bandwidth of the post-perihelion data included essentially all of the ν_3 band. Fifteen spectral lines of water were detected, including three previously undetected lines in the ν_3 band ($\lambda \simeq 2.65$ μm) and three lines in the (011–010) band. The observed relative intensities illustrate that water in the cometary coma is rotationally relaxed, as predicted by recent nonthermal-equilibrium models[2,3]. An *ortho/para* ratio of ~3 is consistent with both pre- and post-perihelion data. The water spatial brightness profile can be fit by that of a parent molecule, and a significant sunward–tailward brightness asymmetry was observed which suggests ejection primarily into the sunlit hemisphere. Comet Halley exhibited marked temporal activity on timescales as short as 2 h. The largest water production rate measured was 2.3×10^{30} mol s^{-1} on 24:18 UT March 1986, which is about an order of magnitude (or more) larger than the production rates we measured pre-perihelion at the same heliocentric distance (~1 AU), indicating a large pre- to post-perihelion asymmetry in gas production.

Spectra were obtained on 20, 22, 24 and 26 March 1986 using the University of Arizona high-resolution Fourier-transform spectrometer (FTS) The heliocentric distance (R) was ~1 AU for all dates while the geocentric distance (Δ) decreased from 0 8 to 0 6 AU during our investigation Observations from airplane altitudes (~12 km) and large cometary geocentric radial velocities ($\Delta \simeq -43$ km s^{-1} for the period of our observations) were required to obtain good atmospheric transmittance at the positions of cometary H_2O emission lines Comet Halley was in the airplane's observing 'window' for up to 2 5 h during our flights For monitoring the temporal variability of the comet, at least one spectrum was taken each flight through a narrow filter ($\Delta\lambda \leq 0\ 06$ μm) using the relatively modest spectral resolution of ~1 4 cm^{-1} Approximately 15 min were required to obtain these data when our aperture (~41 arc s diameter) was centred on the nucleus Spatial mapping was accomplished using the same instrumental setup, but with the aperture offset by known amounts from the centre of brightness of the comet On 26 March we observed comet Halley at ~0 75 cm^{-1} resolution through a much wider filter ($1\ 9 \leq \lambda \leq 2\ 7$ μm), which transmits essentially the entire ν_3 band of H_2O On 22 March we used the narrow filter at ~0 014 cm^{-1} resolution ($\Delta v = 1\ 2$ km s^{-1}), but these data will be discussed elsewhere[5]

Figure 1 shows the spectra acquired through our wide bandpass filter Fig 1*a* is a spectrum of the Moon showing the atmospheric transmittance, and Fig 1*b* is the cometary emission spectrum Figure 1*a* is characteristic of an H_2O absorption line spectrum in thermal equilibrium at ~216 K and shows significant absorption for rotational levels up to $J=8$ ($J\equiv$total angular momentum quantum number) In contrast, the cometary spectrum is heavily weighted towards low J, with all detectable emission originating from transitions having $J\leq3$ as the lower level

The intensities for all observed H_2O lines are given in Table 1 At least six new H_2O lines were detected post-perihelion that were not observed pre-perihelion Most of these were not observed pre-perihelion simply because the comet had significantly stronger emissions post-perihelion (that is, Q_{H_2O} was

---

[||] Present address Space Telescope Science Institute, Homewood Campus, Baltimore, Maryland 21218, USA.

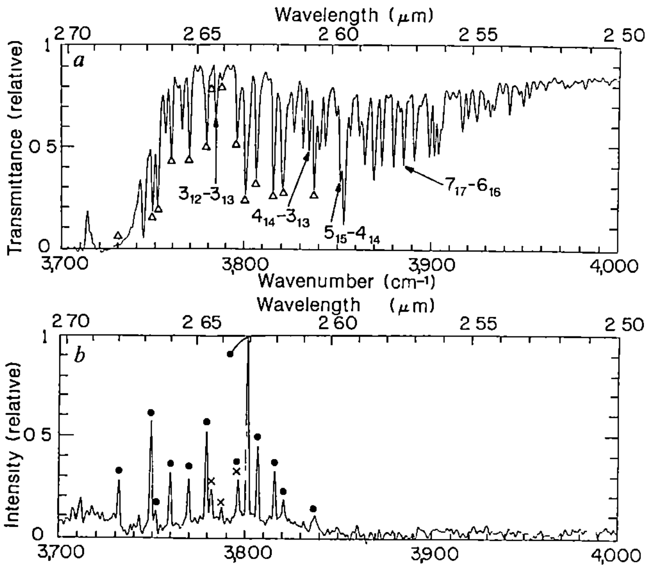


Fig. 1 *a*, A spectrum of the Moon. *b*, The spectrum of comet Halley. Both spectra, taken on 26 March 1986, are apodized to a resolution of 1.5 cm^{-1}, and are uncorrected for atmospheric transmission and instrumental response. For the Halley spectrum, R = 1.096 AU and Δ = 0.646 AU. The absorption lines in the lunar spectrum are due to H_2O and CO_2 in the Earth's atmosphere. Terrestrial H_2O lines corresponding to observed cometary lines are indicated (Δ). The relative strengths of the H_2O lines are characteristic of thermal equilibrium with $T \sim 216$ K. Four H_2O lines whose lower states are the upper states for rotational transitions which have been suggested as probes of H_2O in comets are also identified. These lines are not present in the cometary spectrum. The cometary spectrum is due primarily to molecular line emission produced in solar infrared fluorescence (a little cometary continuum may also have been detected). Most lines are in the (001-000) vibrational band (●), but three lines can be seen in the (011-101) band (×). Note the contrast between the LTE (low thermal emission) terrestrial transmittance spectrum (*a*), which shows significant H_2O absorption for $J \leq 8$, and the non-LTE cometary spectrum (*b*) which sharply cuts off with $J \leq 3$. ($J \equiv$ lower state total angular momentum quantum number)

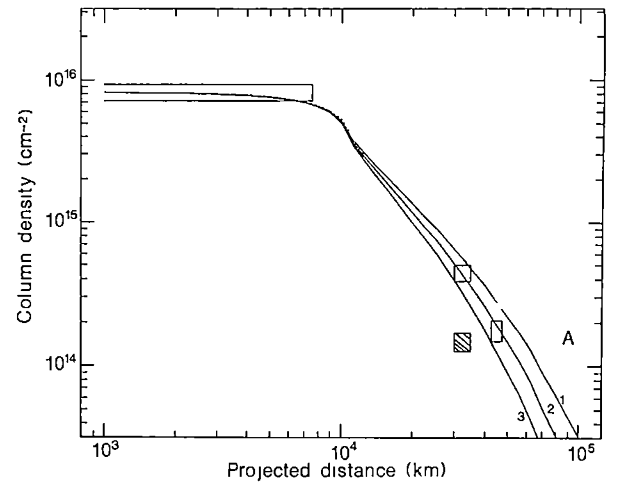


Fig. 2 Assuming that the H_2O emission in the ν_3 band is proportional to the H_2O column density, the boxes represent the observed H_2O spatial brightness distribution. The width of each box is the estimated pointing uncertainty for each observation and the height is the error in the measured relative intensities. The largest component of this error (~10%) is the uncertainty in deriving the relative ν_3 band emission for the nucleus pointings versus the offset pointings (the excitation changes with position in the coma). Two of the offset pointings were towards the Sun while the hatch box represents a tailward pointing. A significant sunward-tailward brightness asymmetry was measured, so only the sunward points are used to determine the H_2O scale-length. Although Q_{H_2O} was varying significantly over the time required for molecules to flow from the nucleus to the offset points, we have ignored this effect. Comparison of KAO and IUE data suggest that Q_{H_2O} was increasing before our observations, which implies that the scale-length derived from our data is a lower limit to the actual value. The curves represent steady-state radial outflow models which have been convolved with a 41 arc s diameter aperture. Scale-lengths: model A, 8.2×10^4 km, model 1, 4.4×10^4 km, model 2, 3.2×10^4 km, model 3, 2.4×10^4 km.

much larger) and/or the lines were not transmitted by our pre-perihelion filter. There are three lines in our spectra which belong to the (011-010) 'hot' band of H_2O. Our calculations of the *g*-factors for this band yield predicted intensities which match the observed values. One of these lines is so strong that it should have been detected in the pre-perihelion spectrum given similar excitation conditions pre- and post-perihelion. The absence of this line pre-perihelion remains to be explained.

The lines originating from levels other than the lowest *ortho* and *para* states (1_{01} and 0_{00}, respectively) have become relatively stronger post-perihelion. This behaviour seems consistent with the hypothesis that larger post-perihelion water production rates cause the pure rotational transitions (which allow relaxation of the H_2O molecules to the lowest states) to become extremely optically thick, thereby 'pumping up' the higher J levels. We note, however, that this mechanism apparently still does not significantly populate levels with $J > 3$. In fact, assuming that the 4_{04}-3_{03} line is pumped only from 3_{03}, we use the measured post-perihelion intensity of this line to derive a relative (to the total) population of 4.3 ± 1% in the 3_{03} level, confirming our previous conclusion[1] that the population in 3_{03} must be small.

Figure 1*a* shows the positions of several strong terrestrial H_2O lines in the ν_3 band whose lower states are the upper states of some rotational transitions which have been suggested as probes of cometary H_2O (6_{16}-5_{23} at 22 GHz, 3_{13}-2_{20} at 183 GHz, and 4_{14}-3_{21} at 380 GHz). The absence of cometary infrared fluorescence (Fig. 1*b*) in these lines (3_{12}-3_{13}, 4_{14}-3_{13}, 5_{15}-4_{14} and 7_{17}-6_{16}) suggests that the rotational levels 3_{13}, 4_{14} and 6_{16} are not significantly populated. This may explain some past failures to detect H_2O rotational lines in comets. We observe ν_3 band transitions which pump the levels 3_{13} and 4_{14} of the ground vibrational state, but the equilibrium populations achieved by the pumps are insufficient to produce detectable infrared fluorescence from those levels. If the rotational transitions 3_{13}-2_{02} and 4_{14}-3_{03} are optically thin, then the emission in two of the above rotational transitions can be estimated directly from the infrared emission rate. We calculate aperture-averaged intensities of 2.3×10^{-16} W cm^{-2} sr^{-1} for the 183 GHz line and 2.4×10^{-16} W cm^{-2} sr^{-1} for the 380 GHz line. Using a Haser model, these intensities can be used to predict the fluxes in a diffraction-limited (for a 0.9-m telescope) aperture. We obtain 28 Jy and 10 Jy for the 183 GHz and 380 GHz lines, respectively, which are below detection limits. However, these predicted fluxes should be considered as lower limits to the actual values because optically thin conditions were assumed and other possible pumping mechanisms were not considered.

The ratio of *ortho*-H_2O to *para*-H_2O in comets may be indicative of the temperature of the cometary ice[1]. Our post-perihelion data give an *ortho*/*para* ratio of 3.23 ± 0.37, compared with our revised pre-perihelion value of 2.73 ± 0.17 (the pre-perihelion value has been revised slightly due to a reassessment of the atmospheric transmittances). Within the present $\pm 2\sigma$ statistical error limits, both the pre- and post-perihelion data are in agreement that the *ortho*/*para* ratio in comet Halley is ~3. More precise determination of the atmospheric transmittances is in progress, and should lead to reduced error estimates for the cometary line intensities, and for the *ortho*/*para* ratio.

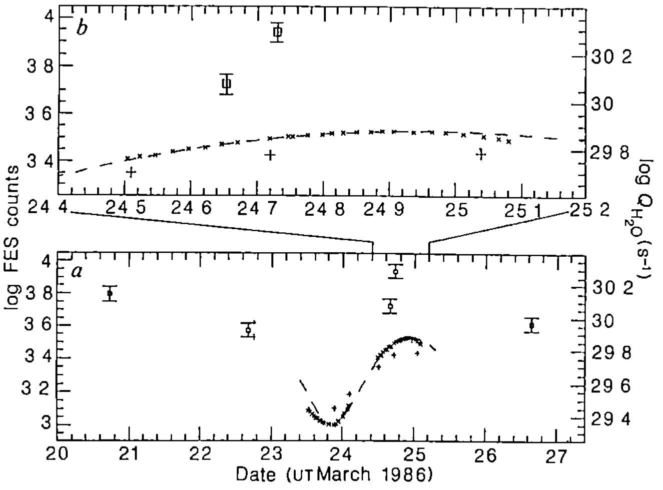


Fig. 3 *a*, Water production rates (using model A parameters) derived for all KAO observations (☐). The error bars are determined solely by our estimate of how pointing inaccuracies can affect our individual observations. The statistical error in the measurements is negligible by comparison. IUE data for 23 and 24 March are also shown. The IUE water production rates (+) are derived from measurements of OH. The integrated visual intensities measured by the IUE fine error sensor (FES) are also given (×). The dashed curve is a sinusoid of period 2.2 days and gives a fairly good match to the FES curve. Propagating this sinusoid forward and backward in time would yield peaks in activity at the positions of the dotted vertical lines. *b*, An expansion of *a* during the time of simulaneous KAO and IUE observations. The temporal variability of the KAO data seems to be substantially larger than that of the IUE data for the time around 24.17 March UT.

Table 1 Intensities of H_2O lines

Transition*	Rest frequency† (cm^{-1})	Intensities‡
000-101 (001-000)	3,732.135	154±78
111-110 (001-000)	3,749.330	114±23
220-221 (001-000)	3,752.213	19±4
110-111 (001-000)	3,759.845	38±6
211-212 (001-000)	3,769.890	37±8
101-000 (001-000)	3,779.494	42±3
202-101 (011-010)	3,782.179	14±1
211-110 (011-010)	3,787.666	6±1
313-212 (011-010)	3,796.083	7±1
212-111 (001-000)	3,796.440	18±1
202-101 (001-000)	3,801.420	100±5
211-110 (001-000)	3,807.015	44±3
313-212 (001-000)	3,816.092	30±3
303-202 (001-000)	3,820.739	12±2
404-303 (001-000)	3,837.870	12±3

* The quantum numbers of the upper and lower states $J'K'_aK'_c$-$J''K''_aK''_c$, and the vibrational band designation ($v'_1v'_2v'_3$-$v''_1v''_2v''_3$) are given.

† From ref. 12.

‡ Relative intensities are listed. The integrated line intensity for the 202-101 (001-000) transition was 9.24 (±0.09)×10^{-10} W cm^{-2} sr^{-1} on 24.18 UT March 1986. The estimated ν_3 band intensity in this case was 1.18 (±0.11)×10^{-8} W cm^{-2} sr^{-1}. The quoted errors are 1σ statistical errors plus our estimate of the uncertainties in the atmospheric transmittances.

Spatial information on the H_2O emission is required for determining the H_2O scale-length. On 24 March 1986 we made observations with sunward and tailward (approximately) offsets of 65 arc s (3.3×10^4 km at the comet) relative to the nucleus, and bracketed these by two exposures centred on the nucleus. On 26 March we made an observation centred on the nucleus and one offset by 97 arc s sunward.

As illustrated in Fig. 2, a significant sunward-tailward brightness asymmetry is observed, which suggests ejection of H_2O primarily from the sunlit hemisphere. The Giotto multi-colour camera[6] pictures of Halley showed an asymmetry in nucleus activity directly, and our data demonstrate that the asymmetry persists far from the nucleus. We will ignore this asymmetry and derive an H_2O scale-length using only the sunward brightness profile and a spherically symmetric model.

Assuming that the measured brightnesses are linearly proportional to the H_2O column density, we show predicted brightness profiles for several scale-lengths in Fig. 2. The observed H_2O spatial brightness distribution is consistent with that of a parent molecule. The scale-length which best matches our data is ~3.2×10^4 km (model 2, R = 1 AU). Taking 1 km s^{-1} as the nominal outflow velocity implies an H_2O lifetime which is 2.8 times shorter than the photochemical value at solar maximum[7]. Our very high resolution data[5] are consistent with an H_2O velocity of ~1.4 km s^{-1}, indicating that the H_2O lifetime could be shorter still.

The scale-length derived from the Giotto neutral mass spectrometer[8] is given approximately by model 1 and is marginally consistent with our data. However, our data are not consistent with model A, which is based on the calculated photochemical lifetime. The short H_2O lifetime implied by both the Giotto and KAO observations is surprising because the Halley apparition occurred near the time of solar minimum when the H_2O lifetime should have its largest value (even larger than that of model A).

Temporal variation in Q_{H_2O} may explain why our data give short apparent scale-lengths. For a pointing centred on the nucleus, ~70% of the observed H_2O molecules are in a sphere whose radius is equal to the projected size of our aperture radius at the comet (1.04×10^4 km). If the molecules are moving at ~1 km s^{-1}, then it takes ~2.9 h for an H_2O molecule to move from the nucleus to the edge of this sphere. Thus, the nuclear pointings mainly represent an integral of Q_{H_2O} over the previous 3 h. As ~9 h and ~13 h are required for H_2O molecules to reach the two offset positions. These pointings represent an integral of Q_{H_2O} over a period beginning much earlier than the time sampled by the nucleus pointing. Analysis of KAO and International Ultraviolet Explorer (IUE) data[9] suggests that Q_{H_2O} was rising during the 12 h preceding our observations. If this is true, then the scale-length derived from our data will be a lower limit to the actual value.

Using a nominal H_2O velocity of 1 km s^{-1}, and the scale-length given by model 2 of Fig. 2, we obtain a water production rate of 2.3×10^{30} mol s^{-1} for the observation with the largest H_2O emission (24.18 UT March). Below, we compare our data with water production rates derived from IUE measurements of OH, which also assume isotropic production and use the H_2O parameters of model A (v = 1 km s^{-1}). For this model, the KAO data imply a production rate of 1.9×10^{30} mol s^{-1}.

The model A water production rates derived from our observations are plotted in Fig. 3. Simultaneous IUE observations[9] were made on 24 March, and the derived H_2O production rates are plotted in Fig. 3, along with the count rate from the IUE fine error sensor (FES, this is an acquisition camera being employed essentially as a photometer during these Halley observations).

The FES data display an apparent harmonicity with a period which is tantalizingly close to the value suggested by others[10] as the rotation period of the comet (2.2 days). Assuming that there is a fundamental period of 2.2 days, we have indicated by the dotted vertical lines in Fig. 3 where the peaks in cometary activity should occur. The post-perihelion KAO data show clearly that the cometary activity cannot be described so simply, a peak in activity would be expected on 22 March, while the data show no evidence for this. Similar departures from simple periodicity were observed for IUE observations made during the week of the spacecraft encounters[11].

All the data illustrate the extremely dynamical behaviour of comet Halley, but the KAO H_2O data show the largest variability At a time when the H_2O brightness increased by ~64%, the FES counts and OH brightness increased by only ≤6% (see Fig 3*b*). These results indicate that the H_2O line brightnesses are very sensitive remote indicators of nucleus activity

The post-perihelion KAO production rates seem to be systematically larger than those derived from the IUE measurements (a factor of ~2) However, the discrepancy between the KAO and IUE numbers may be due in part to absolute calibration uncertainties Based on a comparison of the signal-to-noise ratio achieved on our data, the nominal instrument sensitivity, and the absolute fluxes produced in our calibration process, we expect that the KAO absolute intensities (from which we derive Q_{H_2O}) should be accurate to within ±50% Because the IUE absolute calibration uncertainty is ±10%, the peak water production rates derived from the IUE measurements are actually consistent (at the 1σ level) with the average KAO value

Besides absolute calibration difficulties, several other factors must be considered when comparing the KAO and IUE results Both the KAO and IUE data have been reduced assuming that all cometary emissions are optically thin The aperture-averaged optical depths for the strongest lines are ~1 for both the KAO and IUE observations, so this assumption is not valid However, a more important consideration is the effect of temporal activity on the derived Q_{H_2O} Our data indicate that Q_{H_2O} derived from the KAO observations is more sensitive to temporal activity than Q_{H_2O} derived from the IUE observations Until time-dependent effects are treated properly, strict comparisons of water production rates derived from the two sets of data are not valid Similar comments should apply to water production rates derived from radio OH observations

Finally, we note that both our pre- and post-perihelion KAO observations were made when the comet was at a heliocentric distance of ~1 AU A significant pre- to post-perihelion asymmetry in the water production rate is observed Our largest post-perihelion rate is a factor of ~10 larger than that derived for 24 December 1985 (this was near the time of an outburst in cometary activity), and a factor of ~25 larger than the value derived for 22 December 1985 In contrast, the water production rates derived from the IUE observations[9] of Halley near 1 AU suggest that there is virtually no pre- to post-perihelion asymmetry Because only relative measurements using the same instrument are important in determining the pre- to post-perihelion asymmetry, the discrepancy cannot be attributed to absolute calibration uncertainties Different sensitivities of the infrared and ultraviolet data to temporal activity may account for the discrepancy, but, failing this, the relationship between OH and H_2O in the cometary coma may require reinterpretation

We thank Lou Haughney and the staff of the Kuiper Airborne Observatory for excellent support The help of Mike Williams was crucial for the smooth operation of the spectrometer We thank Paul Butterworth and William Jackson for performing the IUE OH observations on 24-25 March and for permission to use these data Conversations with Paul Feldman contributed valuable insights into the effects of Halley's temporal variability Financial support from NASA grants NAG 2-356 to The Johns Hopkins University and NAG 2-206 to the Lunar and Planetary Laboratory (The University of Arizona), and from NASA RTOP's 188-41-55, 196-41-54 and 352-02-03 to the Goddard Space Flight Center, is gratefully acknowledged

Received 11 July, accepted 15 September 1986

1 Mumma, M J, Weaver, H A, Larson, H P, Davis, D S & Williams, M *Science* **232**, 1523-1528 (1986)
2 Weaver, H A & Mumma, M J *Astrophys J* **276**, 782-797 (1984)
3 Crovisier, J *Astr Astrophys* **130**, 361-372 (1984)
4 Davis, D S, Larson, H P, Williams, M, Michel, G & Connes, P *Appl Opt* **19**, 4138-4155 (1980)
5 Larson, H P, Davis, D S, Mumma, M J & Weaver, H A *Astrophys J* (in the press)
6 Keller H U *et al Nature* **321**, 320-326 (1986)
7 Huebner, W F & Carpenter, C W *Los Alamos Sci Lab Rep*, LA-8085-MS (1979)
8 Krankowsky, D *et al Nature* **321**, 326-329 (1986)
9 Feldman, P D *et al* in *New Insights in Astrophysics, 8 years of UV Astronomy with IUE* (ed Rolf, E J) *ESA Spec Publ* **263** (1986)
10 Sekanina, Z & Larson, S M *Nature* **321**, 357-361 (1986)
11 Festou, M C *et al. Nature* **321**, 361-363 (1986)
12 Camy-Peyret, C, Flaud, J M, Guelachvili, G & Amiot, C *Molec Phys* **26**, 825-855 (1973)

Identification of [Fe III] in the solar ultraviolet spectrum

C. Jordan*, J.-D. F. Bartoe†, G. E. Brueckner† & K. P. Dere†

* Department of Theoretical Physics, 1 Keble Road,
Oxford OX1 3NP, UK
† Naval Research Laboratory, Washington D C 20375, USA

Solar ultraviolet spectra have been obtained with high spectral and spatial resolution using the Naval Research Laboratory's high-resolution telescope and spectrograph[1] (HRTS) flown on rockets since 1975 and most recently on the Spacelab 2 Shuttle flight. Because the solar spectrum between ~1,170 and 1,719 Å has been well observed for some years, few lines of substantial intensity remain unidentified. (See for example the recent compilation by Sandlin *et al.*[2]). The longest exposures during the first rocket flight of the HRTS, which obtained spectra with a spatial resolution of ~1 arc s along the slit, of length ~1 solar radius, and a spectral resolution of ~0.05 Å, showed a number of weak emission lines at the solar limb, in particular between 1,570 and 1,600 Å (see Plates 16 and 17 in ref. 3), that could not readily be identified. Improved observations of these lines were obtained during the flight of HRTS on the Spacelab 2 Shuttle flight in July–August 1985, by making longer exposures (60, 100 and 250 s) The scope of the data obtained has been described[4]. We have now identified the emission lines concerned as forbidden (electric quadrupole and magnetic dipole) transitions in Fe III, the first detection of these particular transitions in any source. They must now be considered potential candidates for previously unidentified lines in other low-density ($N_e \leq 10^{10}$ cm^{-3}) astrophysical sources.

The wavelengths of the emission lines are given in Table 1 Three groups of lines occur, the strongest between 1,575 and 1,598 Å, the others between 1,503 and 1,518 Å and between 1,431 and 1,452 Å The wavelengths were determined on a scale established from nearby lines of neutral atoms For the group around 1,575 to 1,598 Å these were lines of Si I (ref 5) The wavelength accuracy is estimated as ±0 02 Å for lines in this group and ±0 04 Å for the weaker lines Where Sandlin *et al*[2] include the lines their wavelengths agree with those listed to within these limits Several of the lines under discussion are at wavelengths close to those of lines that appear on the disk or in plage and sunspot spectra from the HRTS I flight However, their behaviour at the limb is quite distinct and unlike the typical appearance of the species, such as neutral atoms, with which they could be blended The lines are observed only at and above the solar limb and are not detectable more than ~20 arc s inside the limb It was this unusual feature that first attracted attention In the Spacelab 2 spectra the lines are most clearly observed when the spectrograph slit was set to pass radially through the north solar polar coronal hole They then extend farther above the limb then in 'quiet' regions Figure 1 shows one such spectrum (HRTS Spacelab plate 266, 60 s exposure) Although weak, the lines extend farther above the limb than nearby stronger lines of C I and Fe II, indicating a degree of ionization greater than that of Fe II This conclusion is supported by a series of spectra, also obtained during the Spacelab mission, where the spectrograph slit was set tangential to the limb, centred close

Table 1 Transitions of [Fe III] in solar limb spectra

Observed* λ(Å)	Predicted λ(Å)	Transition J_{lower}–J_{upper}		Notes
		a^5D–a^5G		
1,575.62	1,575.64	4	5	weak
1,576.68	1,576.66	4	6	strongest line
1,585.99	1,586.03	3	4	blend on disk with Fe II (see ref. 2)
1,586.52	1,586.54	3	5	
1,593.50	1,593.50 1,593.48	2 2	3 2	
1,593.67	1,593.68	2	4	blend on disk with Fe II and Ni II (see ref. 2)
1,598.42	1,598.41	1 1	3 2	blend on disk with Fe II (see ref. 2)
		a^5D–a^5P		
1,503.23	1,503.24	4	2	blend on disk
1,504.53	1,504.56	4	3	strongest in multiplet
1,511.58	1,511.59	3	1	blend in HRTS I spot
1,514.48	1,514.50	3	3	
1,518.51	1,518.54	2	1	v. faint
		a^5D–b^5D		
1,431.93	1,431.91 1,431.89	4 4	3 2	
1,434.77	1,434.81	4	4	
1,443.84	1,443.84	3	4	blend on disk
1,447.20	1,447.20 1,447.22	2 2	2 3	blend on disk
1,451.27	1,451.27	1	2	v. faint, blend on disk
1,452.32	1,452.31	1	1	blend on disk

* Accuracy estimated as ±0.02 Å for lines in a^5D–a^5G, and ±0.04 Å for other lines.

to the boundary of the north polar coronal hole. The slit was stepped from inside the limb (at −5 arc s) to above the limb (up to +17 arc s) in increments of 1 arc s (up to +5 arc s) and then 2 arc s. Although the absolute pointing accuracy has yet to be determined, the stigmatic spectra obtained show clearly the reduction in the extent of the chromospheric continuum and successive reduction in the length of lines from different stages of ionization as the slit was moved to greater heights. The Si I recombination continuum ($\lambda \leq 1,520$ Å) and lines of Si I, formed in the solar chromosphere, are restricted to the lowest heights. Higher ionization or excitation lines (for example those of C IV at 1,548 and 1,550 Å and He II at 1,640 Å) extend farthest out into the corona. The lines of Fe II and Si II are intermediate in their behaviour. The lines of [Fe III] can be seen only near the edges of the spectra, indicating strong limb brightening. They extend farther than Si II lines of comparable strength, are broader than the Si I lines and extend at least as far as much stronger lines of Fe II.

The strong limb brightening of the lines shows that they are optically thin transitions, behaviour shared with semi-forbidden electric dipole lines, for example the ^3S–^5S lines of O III at 1,661 and 1,666 Å. This indicates transitions either in an abundant species of low oscillator strength or in a species of low abundance. The spatial behaviour suggests at least a second or third stage of ionization. The available spatial information is consistent with the present proposal of lines of [Fe III].

The term scheme of Fe III[6,7] is illustrated in a partial way in Fig. 2. It is unusual in that states of even parity in the $3d^6$ and $3d^54s$ configurations extend to 80,000 cm^{-1} above the ground state before the first term of odd parity occurs. Thus permitted resonance transitions (to the ground a^5D term) do not begin until wavelengths ≤1,132 Å. Permitted transitions at longer wavelengths occur only to even parity states well above the ground state. The solar lines are transitions within the multiplets $3d^6\ a^5D$ to $3d^5\ (a^4G)\ 4s\ a^5G$, to $3d^5\ (a^4P)\ 4s\ a^5P$ and to $3d^5\ (a^4D)\ 4s\ b^5D$ and include both electric quadrupole and magnetic dipole transitions. (Only one other quintet term, $3d^3(^6S)4s\ ^5S$, lies between a^5G and the ground a^5D term. Transitions between a^5D and a^5S would lie between 2,439 and 2,502 Å, outside the range of the HRTS instrument.) Fe III is a well-known emitter of forbidden lines in the optical spectra of gaseous nebulae and peculiar stars (see for instance refs 8 and 9). However, this is the first identification of [Fe III] in the vacuum ultraviolet region. Table 1 shows the observed and predicted wavelengths. The strongest line is in the a^5D–a^5G multiplet between $J=4$ and $J=6$, as one might expect for a quadrupole transition where $\Delta L=2$ (see ref. 8).

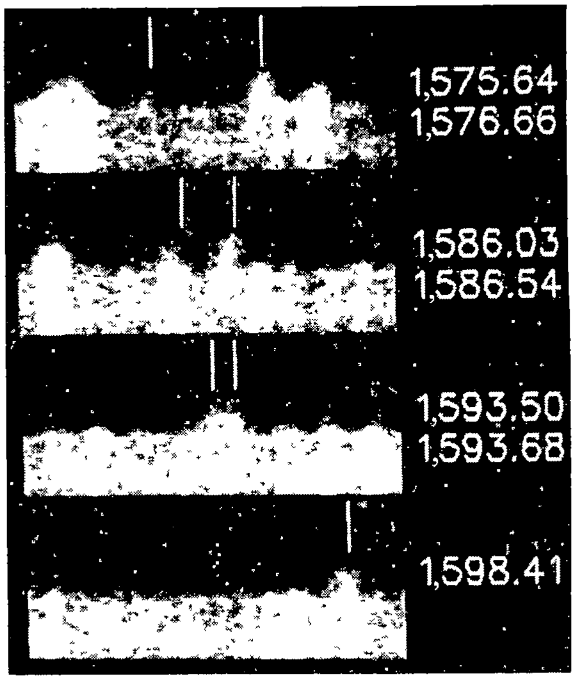


Fig. 1 Sections of a solar limb spectrum obtained with the HRTS instrument on the Spacelab 2 Shuttle flight. HRTS plate 266, a 60 s exposure with the slit radial, crossing the north polar coronal hole. The spectra as illustrated extend about 25 arc s on to the Sun. The wavelengths are given in ångstroms. The wavelength increases to the right.

The intensity calibration of the Spacelab 2 spectra is still in progress, but the order of magnitude of the line intensities can be found from previous spectra[2]. The strongest [Fe III] lines have intensities at the limb ~15 erg cm^{-2} s^{-1} sr^{-1}, about an order of magnitude weaker than the strongest lines of Fe II and a factor of ~100 weaker than the C IV resonance lines.

Until transition probabilities and collision cross-sections are available it is not possible to make definitive statements concerning the excitation of the lines. At present these atomic data have been calculated only for optical forbidden transitions[8,9] or for electric dipole transitions[10]. The [Fe III] lines are likely to be formed at $T_e \approx 20$–30 kK according to ion-balance calculations[11]. The electron density in a coronal hole at this temperature is $N_e \approx 10^{10}$ cm^{-3} (see for instance ref. 12). From the calculations by Garstang *et al.*[9], collision strengths, Ω, between ~1 and 0.3 are typical for transitions with $\Delta L=0$, 1 and 2 and $\Delta J=0$, 1 and 2. If a mean value of $\Omega=0.5$ is adopted then it can be shown that collisional de-excitation of the upper levels will exceed spontaneous radiative decay unless $A>40$ s^{-1}. The transitions listed in Table 1 can occur either by electric quadrupole radiation ($\Delta J=2$) or by both magnetic dipole and electric quadrupole radiation ($\Delta J=0$, 1). The largest transition prob-

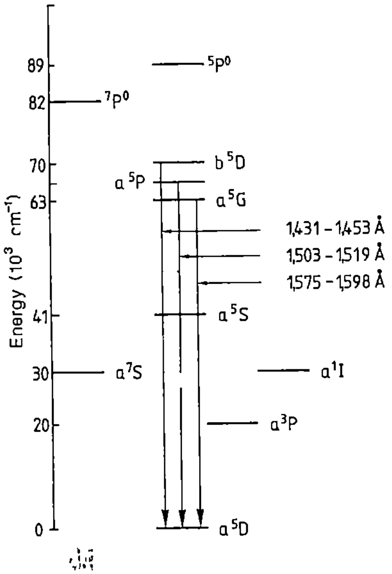


Fig. 2 A partial term scheme for Fe III, showing the observed transitions, some lower-lying terms of the same even parity and the first two terms of odd parity Term energies are given in cm^{-1}

abilities found by Garstang[8] for lower-lying levels are in each case ~1 Thus it is likely that the upper levels have populations close to their Boltzmann values with collisional de-excitation rates exceeding radiative decay rates Then the ratio of the [Fe III] line fluxes to those of permitted transitions would be sensitive to the electron density Detailed calculations of the atomic data are required to establish the collisional-radiative regime for the individual lines The upper levels may attain only a pseudo-Boltzmann population if collisional excitation to higher states of odd parity exceeds the rate for collisional de-excitation to lower levels In any case the new [Fe III] identifications will provide more information on the structure of the solar chromosphere-corona transition region

If the a^{5}G, a^{5}P and b^{5}D levels are collisionally de-excited in the solar atmosphere, they could become stronger, relative to permitted transitions of species of similar excitation, in astrophysical sources of lower electron density For this reason their presence is being investigated in such sources, including the Seyfert galaxy NGC 4151, which has unidentified emission features around 1,575, 1,581 and 1,518 Å (refs 13 and 14) [Fe III] emission is observed in the optical spectrum of NGC 4151[15,16] The relative intensities of the quintet transitions in NGC 4151 and other sources cannot be predicted until collision cross-sections and transition probabilities are known, and these are urgently required to establish whether or not the [Fe III] lines are of wider astrophysical significance

Received 3 September, accepted 23 October 1986

1 Bartoe, J-D F, Brueckner, G E, Purcell, J D & Tousey, R. *Appl. Opt.* **16**, 879-885 (1977)
2 Sandlin, G D Bartoe, J-D F, Brueckner, G E, Tousey, R & VanHoosier, M E *Astrophys J suppl. Ser* **61**, 801-898 (1986)
3 Jordan, C, Brueckner, G E, Bartoe, J-D F, Sandlin, G D & VanHoosier, M E *Astrophys J* **226**, 687-697 (1978)
4 Bartoe, J-D F *et al.* in *AIAA 24th Aeropsace Sciences Meeting, Reno*, AIAA-86-0225 (in the press)
5 Moore, C E in *Selected Tables of Atomic Spectra Atomic Energy Levels and Multiplet Tables, Si L* NSRDS-NBS-3, Sect 2 (1967)
6 Edlen, B & Swings, P *Astrophys J* **95**, 532-554 (1942)
7 Reader, J & Sugar, J *J phys chem. Ref Data* **4**, 353-440 (1975)
8 Garstang, R. H. *Mon. Not. R. astr Soc.* **117**, 393-405 (1957)
9 Garstang, R. H., Robb, W D & Rountree, S P *Astrophys J* **222**, 384-397 (1978)
10 Abbott, D C *J Phys* **B11**, 3479-3497 (1978)
11 House, L L *Astrophys J suppl. Ser* **8**, 307-328 (1964)
12 Munro, R. H & Withbroe, G L *Astrophys J* **176**, 511-520 (1972)
13 Ulrich, M H *et al. Nature* **313**, 747-751 (1985)
14 Cassatella, A & Jordan, C (in preparation)
15 Boksenberg, A. *et al. Mon Not R. astr Soc.* **173**, 381-386 (1975)
16 Boksenberg A & Penston M V *Mon Not R astr Soc.* **177** 127*P*-131*P* (1976)

A hierarchical $O(N \log N)$ force-calculation algorithm

Josh Barnes & Piet Hut

The Institute for Advanced Study, School of Natural Sciences,
Princeton, New Jersey 08540, USA

Until recently the gravitational N-body problem has been modelled numerically either by direct integration, in which the computation needed increases as N^2, or by an iterative potential method in which the number of operations grows as $N \log N$. Here we describe a novel method of directly calculating the force on N bodies that grows only as $N \log N$: The technique uses a tree-structured hierarchical subdivision of space into cubic cells, each of which is recursively divided into eight subcells whenever more than one particle is found to occupy the same cell. This tree is constructed anew at every time step, avoiding ambiguity and tangling. Advantages over potential-solving codes are: accurate local interactions; freedom from geometrical assumptions and restrictions; and applicability to a wide class of systems, including (proto-)planetary, stellar, galactic and cosmological ones. Advantages over previous hierarchical tree-codes include simplicity and the possibility of rigorous analysis of error. Although we concentrate here on stellar dynamical applications, our techniques of efficiently handling a large number of long-range interactions and concentrating computational effort where most needed have potential applications in other areas of astrophysics as well.

Until recently, the dynamics of a system of self-gravitating bodies (the gravitational N-body problem) has been modelled numerically in two fundamentally different ways The first one, direct N-body integration, involves the computation of all $\frac{1}{2}N(N-1)$ forces between all pairs of particles This allows an accurate description of the dynamical evolution but at a price that grows rapidly for increasing N[1] The second way involves a two-step approach after fitting the global potential field to a special model with a number of free parameters, each particle is propagated in this background field for a short time before the same procedure is reiterated The potential method involves a number of operations that grow only as $N \log N$ Thus calculations can be performed more quickly, but with a loss of accuracy and generality The special nature of each potential-solving code is caused by the need to use some technique that is tuned to the geometry of the problem being considered (such as Fourier transforms or spherical or bispherical harmonics[2])

Recently, some of the advantages of both approaches have been combined by using direct integrations of force while grouping together increasingly large groups of particles at increasingly large distances This corresponds to the way humans interact with neighbouring individuals, further villages and increasingly further and larger states and countries—driven by increasing cost and decreasing need to deal with more removed groups on an individual basis The first implementation of such a hierarchical grouping of interactions was given by Appel[3], who used a tree structure to represent an N-body system, with the particles stored in the leaves of the tree An independent implementation by Jernigan[4] and Porter[5] incorporated regularization of close encounters However, in both codes the logarithmic-growth gain in efficiency comes at the price of introducing additional errors that are hard to analyse because of the arbitrary structure of the tree Nearby particles may be grouped as leaves of nearby branches, but the phase-space flow of realistic self-gravitating systems demands a continuous updating of the tree structure to avoid tangling and unphysical grouping, requiring complicated book-keeping It is not at all clear how to understand and estimate the errors caused by the process of approximating lumps of particles together as single pseudo-particles, because individual lumps can take more or less arbitrary shapes and sizes

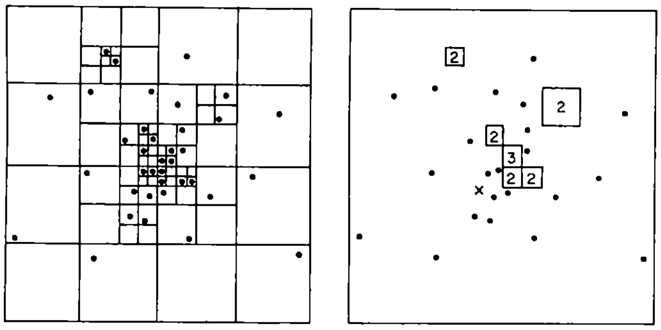


Fig 1 Hierarchical boxing and force calculation, presented for simplicity in two dimensions On the left, a system of particles and the recursive subdivision of system space induced by these particles Our algorithm makes the minimum number of subdivisions necessary to isolate each particle On the right, how the force on particle *x* is calculated Fitted cells contain particles that have been lumped together by our 'opening angle' criterion, each such cell represents a single term in the force summation

We present here a new way of realizing a tree-based force calculation with logarithmic growth of force terms per particle that avoids the tree-tangling complications mentioned above, allows rigorous upper bounds for errors that arise from neglecting internal lump structure, and also offers a well-defined procedure for estimating more typical, average errors The essential ingredients are (1) a virtual cubical division of empty space in (sub)cells with daughter cells having exactly half the length, breadth and width of their parent, (2) the construction of the actual tree of cells from the virtual one by (i) discarding empty subcells, (ii) accepting subcells with one occupant, and (iii) recursively dividing shared occupancies in sub-subcells, and (3) performing this reconstruction *ab initio* at every time step

Given this book-keeping structure, the dynamics are implemented by assigning to every non-empty cell, as well as to higher-order cells containing more than one particle, a (pseudo-)particle that contains the total mass in the cell located at the centre-of-mass of all the particles it contains Any single real particle feels the force of all (pseudo-)particles in the system that represent a cell small enough and far enough to forego the need of further division, thereby screening all its component (pseudo-)particles

A computer program that implements the hierarchical force calculation is available from us upon request It contains less than a thousand lines of C code 150 lines of definitions, 150 lines for tree construction, 100 lines for force calculation and 100 lines for a simple integrator, the remaining lines handle input-output book-keeping

In what follows we summarize some of the more technical details The method we use to compute a force in time of $O(\log N)$ is based on a representation of the mass distribution as a hierarchical tree structure, constructed as follows Begin with an empty cubical cell big enough to contain the system One by one, load particles into this 'root' cell If any two particles fall into the same cell, divide that cell into eight cubical subcells (thus the first such division occurs as soon as the second particle has been loaded in, splitting the system into at least eight pieces) Each divided cell is represented by a data structure that holds information about the subcells it contains a summary of global physical quantities (mass and centre-of-mass position) as well as pointers to the daughter cells, which may be referenced to obtain more detailed information Continue this process of subdividing to as high a level as required When all N particles have been loaded, the system space will have been partitioned up into a number of cubical cells of different sizes, with at most one particle per cell These particle-bearing cells are grouped together into larger cubical cells, which are grouped together into still larger parent cells, and so on down to the root cell, which contains the entire system The average size of a particle-bearing cell is of the order of the interparticle spacing, so the 'height of the tree' (that is, the number of subdivisions required to reach a typical cell, starting at the root), is of $O(\log_2 N^{1/3}) = O(\log N)$, and the time required to construct the tree is of $O(N \log N)$ The final step in constructing the tree is to tag the subdivided cells with the total mass and centre-of-mass position of the particles they contain, by propagating information down the tree from the particles towards the root, this step may also be accomplished in a time of $O(N \log N)$

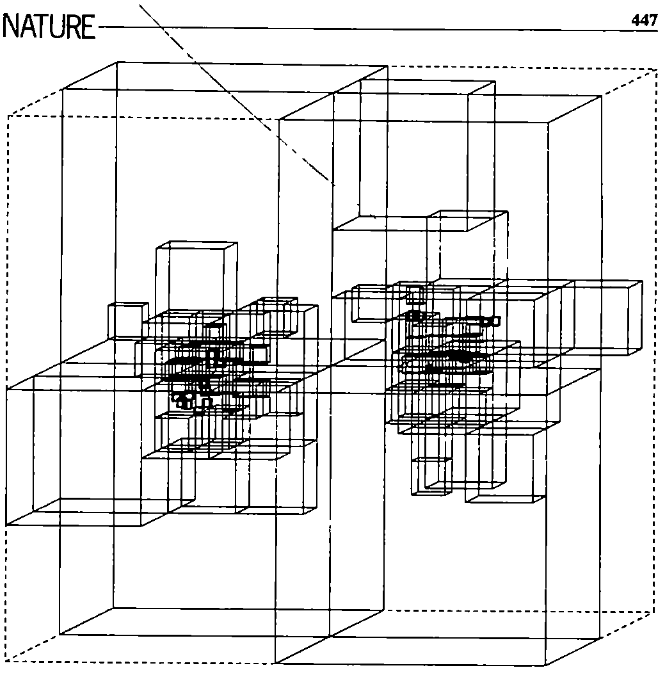


Fig 2 Box structure induced by a three-dimensional particle distribution This example was taken from the early stages of an encounter of two $N = 64$ systems, and shows how the boxing algorithm can accommodate systems with arbitrarily complicated geometry The particle distribution corresponding to a system with 32 times as many members is shown in Fig 3

Having constructed such a tree, the force on any particle p may be approximated by a simple recursive calculation Start at the root cell of the tree, which contains the entire system Let l be the length of the cell currently being processed and D the distance from the cell's centre-of-mass to p If $l/D < \theta$, where θ is a fixed accuracy parameter ~1, then include the interaction between this cell and p in the total being accumulated Otherwise, resolve the current cell into its eight subcells, and recursively examine each one in turn The core of the force calculation routine may be compactly expressed in SCHEME, a dialect of LISP

```
(define (acceleration particle ensemble)
  (cond ((singleton?  ensemble)
         (newton-acceleration particle (the-element ensemble)))
        ((< (/ (diameter ensemble)
               (distance particle (centroid ensemble)))
            theta)
         (newton-acceleration particle (centroid ensemble)))
        (else
         (reduce sum-vector
                 (map (lambda (e) (acceleration particle e))
                      (subdivisions ensemble))))))
```

Note that in LISP, a function with arguments $f(x, y, \quad)$ is written as $(f\,x\,y \quad)$ For example, (newton-acceleration p_1 p_2) calls a function to compute the acceleration of particle p_1 due to p_2 The (cond) form is a three-way conditional, computing the acceleration directly in the first two cases, and by recursion in the final (else) clause Elements of the SCHEME programming language are presented in Abelson *et al*[6] Figure 1

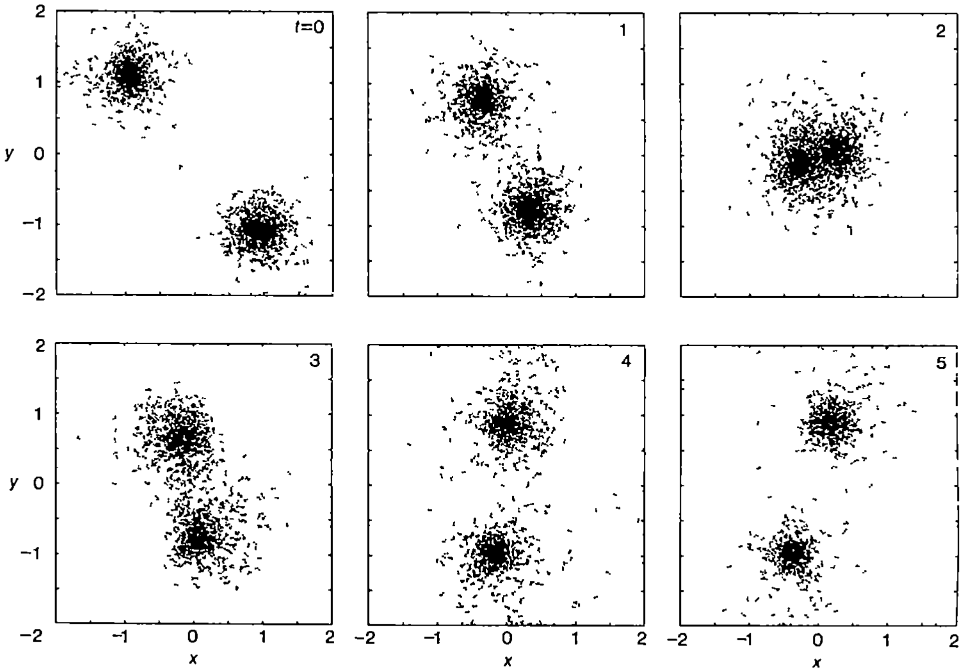


Fig 3 Encounter of two spherical systems, simulated by using our hierarchical acceleration technique in combination with a simple leap-frog integrator The incoming systems were launched on parabolic orbits, but become bound because of dynamical friction Note the striking wakes lagging behind the density centres of the two systems at $t=3$, 4 (in our units the gravitational constant, the mass of each galaxy and the total binding energy of the whole system all equal unity) With a total of $N=4{,}096$ particles in the system and an opening angle criterion of $\theta=1$, the number of two-body interactions computed by our technique is less than 0 1 of the $\frac{1}{2}N/(N-1)$ required by a direct-summation force calculation The calculation took 10 h on a VAX 11/780 (in double precision because of compiler limitations, a single-precision calculation would take half as long) With a time step $\Delta t = 0\ 05$ and softening parameter $\varepsilon = 0\ 025$, energy was conserved to ~1%

illustrates this process for a small number of particles in two dimensions, increasing the number to $10^4 \sim 10^5$ in three dimensions typically increases the number of interactions per particle to only of order 10^2

The number of interactions considered by this procedure in computing the force on p is of order log N for large N Suppose the mass distribution is homogeneous within the root cell Increasing the total number of particles eightfold is roughly equivalent to adjoining eight similar root cells together The seven new cells not containing p will contribute some 'relatively small' number ΔN_t of additional terms to the force approximation Now the expectation value $\langle \Delta N_t \rangle$ depends on θ, but not on the total number of particles or the size of the system Thus the time required to calculate the force on a particle increases by a constant increment (of $\langle \Delta N_t \rangle$), whereas N increases by a constant factor (of eight) In other words, the time required by the CPU (central processing unit) to compute the force on a single particle is on the scale of $O(\log N)$

A rigorous error analysis of the force-calculation algorithm is possible because our prescription yields a unique, well-characterized tree structure based on up-to-date particle positions Each compound cell that we choose not to subdivide introduces a small error due to quadrupole and higher-order moments of the mass distribution within the cell (the dipole term vanishes when expanding around the centroid) The magnitude of this error may be bounded by a 'worst-case' analysis for which the quadrupole moment is maximized (for example, two lumps placed in opposite corners of the cell), and estimated from an analysis of root-mean-square fluctuations within each cell together with estimates of the coherence time scales for these fluctuations We shall present this analysis in a more detailed paper In practice, forces computed even with an opening angle parameter as large as $\theta = 1$ are still accurate to ~1% with little dependence on N Empirically, we find the force error scales approximately as the −1 5 power of the computing time These errors are only weakly correlated from one time step to the next, resulting in a build-up close to a random walk rather than a steady drift

As a test of our new method, we have written a simple N-body code using our force-calculation scheme with a time-centred leap-frog integrator, in which positions and velocities are alternately advanced A parabolic encounter of two galaxies is initiated at a distance of several galactic radii, leading to a box structure as shown in Fig 2 The results of a 4096-body calculation of such an encounter are shown in Fig 3 This calculation took 10 h of CPU time on a VAX 11/780 with a floating point accelerator

There are several ways in which the code can be made more efficient We are now investigating these, and we shall discuss our results in detail elsewhere We just mention three possible improvements (1) using a higher-order integration scheme such as Aarseth's fourth-order polynomial method rather than our second-order leap-frog method, which will require careful adjustments to avoid glitches caused by discrete differences between the grouping of particles in cells from one time step to the next (for example by multiply covering space in partly overlapping virtual grids), (2) including quadrupole moments in the description of cells as pseudoparticles characterized by the total mass in the cell as located in the centre of mass, (3) introducing individual time steps for particles which undergo strongly changing interactions, which could be accomplished by subsequently halving the time step when needed—thus extending the three-dimensional spatial halving of cells to a four-dimensional space-time division in rectangular subcells

An interesting aspect of our new code is the different emphasis it places both on software and hardware, in comparison with other codes On the hardware side, the hierarchical structure of our code does not lend itself easily to vectorization (although this may well be worth exploring) In contrast, we expect our code to be most useful on computers with highly parallel architectures (with one processor per particle, computer time is reduced approximately by a factor of N) On the software side, the hierarchical decomposition of the problem is best realized by using recursive descriptions Recursive function calls and other general control and data structures are not well supported or clearly represented in FORTRAN This has led us to consider other programming languages such as C, PASCAL and LISP Another advantage offered by these languages is that they permit a clarity of presentation of our ideas, which makes the underlying techniques available to other researchers Of course, if a particular computer has a FORTRAN compiler which is an order of magnitude faster than other compilers, it makes sense to translate a version of our program into FORTRAN, trading clarity and modularity for efficiency

Our application to N-body calculations is only one in a range of possibilities including the calculation of radiation fields (replacing particles with sources) and self-gravitating fluid flow (cell division being governed by the complexity of the local flow pattern) Thus our technique forms a general tool for simultaneously handling a large number of long-range interactions and for concentrating computing resources locally where most needed

We thank John Bahcall, Jeremiah Ostriker and especially Gerald Sussman for interesting discussions Part of this work was supported by the NSF through grant PHY-8217352, P H is an Alfred P Sloan Foundation fellow

Received 16 July, accepted 1 October 1986

1 Aarseth, S J in *Multiple Time Scales* (ed Brackbill, J U & Cohen B I) 377-418 (Academic Press, New York, 1985)
2 Hockney, R W & Eastwood, J W *Computer Simulation using Particles* (McGraw-Hill, New York, 1981)
3 Appel, A *SIAM J Sci. statist Comput* **6**, 85-103 (1985)
4 Jernigan, J G *IAU Symp* **113**, 275-284
5 Porter, D thesis, Physics Dept, Univ California, Berkeley (1985)
6 Abelson, H, Sussman, G J & Sussman, J *Structure and Interpretation of Computer Programs* (MIT Press, Cambridge, Massachusetts, 1985)

The 400-km seismic discontinuity and the proportion of olivine in the Earth's upper mantle

Craig R. Bina & Bernard J. Wood

Department of Geological Sciences, Northwestern University, Evanston, Illinois 60201, USA

The 400-km seismic discontinuity has traditionally been ascribed to the isochemical transformation of α-olivine to the β-modified-spinel structure in a mantle of peridotitic bulk composition[1-6]. It has recently been proposed[7,8] that the observed seismic velocity increase at 400 km depth is too abrupt and too small to result from a phase change in olivine but instead requires that the transition zone be chemically distinct in bulk composition from the uppermost mantle. By requiring phase relations in the Mg_2SiO_4-Fe_2SiO_4 system to be internally consistent thermodynamically, we find that the α–β transition in olivine of mantle $(Mg_{0.9}Fe_{0.1})_2SiO_4$ composition is extremely sharp, occurring over a depth interval (isothermal) of ~6 km. The magnitude of the predicted velocity increase is in agreement with that observed seismically[9,10] if the transition zone is composed of ~60–70% olivine. Thus, our results indicate that seismic velocities across the 400-km discontinuity are consistent with a transition zone of homogeneous peridotitic composition and do not require chemical stratification.

The 400-km seismic discontinuity reflects a change in elastic properties of the mantle and has been attributed to a phase transformation of olivine to a spinel-like structure at high pressures[1] Subsequent work has given rise to a generally accepted model in which the discontinuity is attributed to such an isochemical phase change in a mantle of homogeneous olivine-rich, or peridotitic, composition[2-6] This model has the advantage of simplicity and can be tested experimentally

Recently, it has been suggested[7,8] that a phase transition in olivine would produce a gradual velocity increase over an appreciable depth interval—rather than the abrupt increase observed seismically—and that the magnitude of the increase would be more than twice that actually observed It was proposed that the seismic data require the transition zone to be chemically distinct in bulk composition from the uppermost mantle, with the transition zone consisting of a pyroxene-garnet rich 'piclogite' composition containing either 16%[7] or 30%[8] olivine The 400-km discontinuity is ascribed to either a change in chemical composition from peridotite to underlying piclogite or—if this periodotite/piclogite boundary is referred to shallower depths—to the transformation of pyroxene to a garnet-like structure

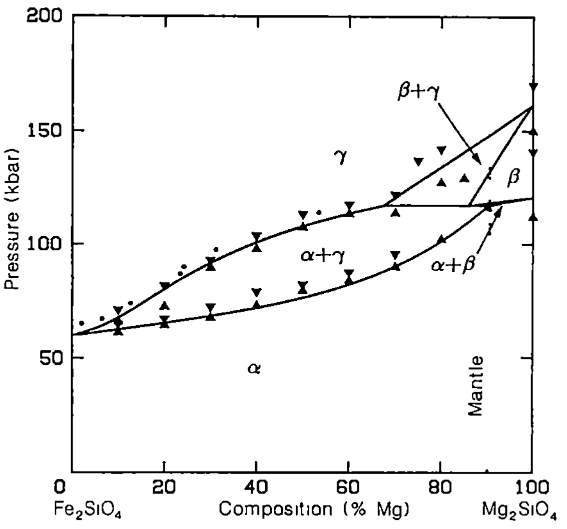


Fig 1 Isothermal pressure–composition diagram showing calculated boundaries for olivine polymorph stability fields at 1,273 K Also shown are experimental data points[12-14,18-20] delimiting the high-pressure stability limits of the low-pressure assemblages (▲), the low-pressure stability limits of the high-pressure assemblages (▼), and the compositions of γ phase which coexist with α phase at the indicated pressures (●) Dashed line shows $(Mg_{0.9}Fe_{0.1})_2SiO_4$ composition

In a previous study[11], we showed that the transformation of pyroxene to a garnet structure would produce a smooth and gradual increase in seismic velocity, rather than the discontinuity observed at 400 km In the present study, we have examined the olivine–spinel phase transitions to determine whether the observed seismic velocity variations may be attributable to such a phase change The α-olivine to β-modified-spinel transition has been commonly represented by a broad 'α+β divariant loop', a region in which both phases coexist in stable equilibrium. If this representation were accurate, the α phase would transform to the β phase in a continuous and gradual manner, and this phase change would not produce a sharp discontinuity in seismic velocity However, the available experimental data (Fig 1) do not constrain the width of this α+β loop, since no high-pressure experiments have yet produced both phases together in equilibrium for olivine of mantle $(Mg_{0.9}Fe_{0.1})_2SiO_4$ composition We have used available thermoelastic and calorimetric data on the olivine polymorphs (α-olivine, β-modified-spinel, and γ-spinel) to constrain the width of the α+β divariant loop By requiring the phase diagram for the Mg_2SiO_4-Fe_2SiO_4 system to be internally consistent thermodynamically, we have attempted to determine the sharpness and magnitude of a seismic discontinuity resulting from a phase change in olivine

If the partial molar free energies of Mg_2SiO_4 and Fe_2SiO_4 components are known as functions of pressure, temperature, and composition, then the boundaries of the stability fields for the various phase assemblages (α, α+β, β, β+γ and so on) can be calculated explicitly To compute the free-energy functions, we require knowledge of the enthalpies, entropies, volumes, and solution activities of the components in the various phases at the pressures, temperatures and compositions of interest We used the available experimentally-measured values of the enthalpies and entropies[5,12-14], heat capacities[15], molar volumes and coefficients of thermal expansion[4], elastic moduli[16], and activity coefficients[17] for the phases and components in question Where measured values were extremely uncertain or

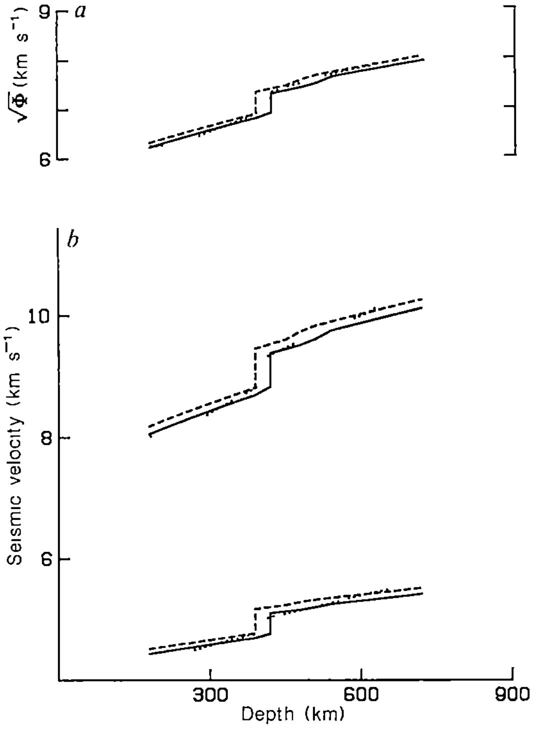


Fig. 2 *a*, Calculated bulk sound velocity ($\sqrt{\Phi}$) as a function of depth for a mantle consisting of pure olivine of $(Mg_{0.9}Fe_{0.1})_2SiO_4$ composition, along 1,700 K (dashed line) and 2,000 K (solid line) isotherms. Composite seismic profile GCA+TNA[9,10] (dotted line) is shown for comparison. *b*, Calculated P- and S-wave velocities (V_P and V_S, respectively) as functions of depth for a mantle consisting of pure olivine of $(Mg_{0.9}Fe_{0.1})_2SiO_4$ composition, along 1,700 K (dashed line) and 2,000 K (solid line) isotherms. Composite seismic profile GCA+TNA (dotted) is shown for comparison.

unavailable (such as for the metastable β-Fe_2SiO_4 component), we estimated them by formally inverting the calculation procedure and requiring the estimated values to reproduce the observed phase relations as constrained by the available synthesis data[12-14,18-20]. (Detailed discussions of the experimental data, the methods of computation and the requirements of thermodynamic consistency are presented elsewhere[21].) Using the Gruneisen thermal state equations[22] and the third-order Birch-Murnaghan isothermal state equation[23] to compute the temperature and pressure dependencies, respectively, of the molar volumes, we computed the boundaries of the stability fields in pressure-composition space for the various assemblages along the Mg_2SiO_4-Fe_2SiO_4 join, as shown in Fig. 1.

The resulting phase diagram is internally consistent thermodynamically and is consistent with the available calorimetric, thermoelastic and synthesis data, within their experimental uncertainties. Note that the $\alpha+\beta$ divariant loop is extremely narrow, ~2 kbar wide, so that the $\alpha \rightarrow \beta$ transition in mantle olivine would be quite sharp, occurring over a depth interval of ~6 km. The width of this transition is extremely robust to uncertainties in the data set, extreme perturbations of the various parameters yield a variation in transition width of only ±1 kbar (~3 km).

Using the method of free-energy minimization[24], we have calculated stable phase assemblages and associated (Voigt-Reuss-Hill averaged[25]) bulk sound velocities ($\sqrt{\Phi}$) as functions of pressure (depth) along 1,700 K and 2,000 K isotherms for a model mantle consisting of pure olivine of $(Mg_{0.9}Fe_{0.1})_2SiO_4$ composition (Fig. 2a). The model bulk sound velocities for a pure olivine mantle agree quite well (within ±0.06 km s^{-1}) with those calculated from the representative high-resolution P-wave and S-wave velocity profiles GCA[9] (Gulf of California) and TNA[10] (Tectonic North America), respectively. (As we are not attempting to model the 670-km discontinuity, we have not included the effects of any higher pressure phases, such as the silicate perovskite, which might contribute to velocity increases at that depth[2].) A sharp jump in velocity occurs at ~400 km depth due to the $\alpha \rightarrow \beta$ transition. A small kink in the velocity profile occurs at ~550 km due to the $\beta \rightarrow \gamma$ transition, but this slight velocity change is barely detectable due to the similarities of the elastic properties of the β and γ phases[16]. The ratio of the magnitude of the observed (GCA+TNA) velocity increase at 400 km (4.9%) to that calculated for the model pure olivine mantle (6.6%) suggests that the mantle in the transition zone contains ~74% olivine.

We have separated our values of bulk sound velocity ($\sqrt{\Phi}$) into P- and S-wave velocities for a model pure olivine $(Mg_{0.9}Fe_{0.1})_2SiO_4$ mantle (Fig. 2*b*). This requires additional assumptions because the temperature and pressure derivatives of the shear moduli of the high-pressure polymorphs are unknown[26]. To perform the decomposition, we first computed velocities by assuming that the behaviour of the shear moduli at high pressures and temperatures follows the same pattern as that of the bulk moduli, namely that the pressure derivatives and the logarithmic temperature derivatives, respectively, are the same[22] for all the phases of a given component. In a second approach, we followed previous workers[27] in extracting P- and S-wave velocities by assuming that Poisson's ratio for our model pure olivine mantle was the same, at corresponding depths, as that observed (GCA+TNA) seismically. These independent approaches yielded virtually identical results and provide support for our conclusions. The magnitudes of the observed (GCA+TNA) velocity increases at 400 km (4.8 and 4.5% in P- and S-wave velocity, respectively), and those calculated for the model pure olivine mantle (7.2 and 8.1%), are indicative of a mantle composed of ~62% olivine (a peridotitic mantle), in excellent agreement with the results of Weidner[6].

Thus, an isochemical phase change from α-olivine to β-modified-spinel in olivine of mantle composition would produce a sharp 400-km seismic discontinuity ~6 km wide. In fact, our computed velocity increase is even sharper than that required by current seismic observations. The actual sharpness of this discontinuity is difficult to resolve seismically[28,29]. Leven[30] suggests that the discontinuity occurs over a depth interval of ~6 km, other studies[31] suggest that the discontinuity is at least 10 km wide. Our transition width of 6(±3) km is computed for isothermal conditions, temperature gradients through the transition could broaden the discontinuity[4] by up to ~7 km.

Thus, seismic velocities above and below a 400 km depth for a model mantle consisting of pure olivine of $(Mg_{0.9}Fe_{0.1})_2SiO_4$ composition are in good agreement with those observed seismically[9,10]. The magnitudes of the abrupt increases in bulk sound velocity and in P- and S-wave velocities at 400 km are consistent with a mantle composed of between 62 and 74% olivine. Since phase transformations from pyroxene to garnet would not produce such a discontinuity[11], a mantle dominated compositionally by olivine appears necessary for the generation of a sharp 400-km discontinuity unless arbitrary changes in bulk chemical composition are invoked. We therefore conclude that seismic observations are consistent with a peridotitic mantle to a depth of at least 650 km and thus do not require the transition region to be chemically distinct from the overlying mantle.

We thank Jim Leven and Hiroshi Sawamoto for providing preprints of papers before publication. We also thank Ray Jeanloz for providing experimental data from the dissertation of Kawada[13], Masaki Akaogi for critical comments, and Don L. Anderson, Jay Bass, Seth Stein and Emile Okal for helpful discussions. This material is based upon work supported under a NSF Graduate Fellowship (to C.R.B.) and NSF grant EAR-8416793 (to B.J.W.).

Received 23 April, accepted 20 August 1986

1 Bernal, J D *Observatory* **59,** 268 (1936)
2 Ringwood, A E *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975)
3 Liu, L in *The Earth Its Origin, Structure and Evolution* (ed McElhinny, M W) 177-202 (Academic, New York, 1979)
4 Jeanloz, R & Thompson, A B *Rev Geophys Space Phys* **21,** 51-74 (1983)
5 Navrotsky, A & Akaogi, M *J geophys Res* **89,** 10135-10140 (1984)
6 Weidner, D J *Geophys Res Lett* **12,** 417-420 (1985)
7 Bass, J D & Anderson, D L *Geophys Res Lett* **11,** 237-240 (1984)
8 Anderson, D L & Bass, J D *Nature* **320,** 321-328 (1986)
9 Walck, M C *Geophys R R astr Soc* **76,** 697-723 (1984)
10 Grand, S P & Helmberger, D V *Geophys J R ast Soc* **76,** 399-438 (1984)
11 Bina, C R & Wood, B J *Geophys Res Lett* **11,** 955-958 (1984)
12 Suito, K in *High Pressure Research—Applications to Geophysics* (eds Manghnani, M H & Akimoto, S) 255-266 (Academic, New York, 1977)
13 Kawada, K thesis, Univ Tokyo (1977)
14 Sawamoto, H *Phys Chem Miner* **13,** 1-10 (1986)
15 Watanabe, H in *High Pressure Research in Geophysics* (eds Akimoto, S & Manghnani, M H) 441-464 (Center for Academic Publications, Tokyo, 1982)
16 Weidner, D J, Sawamoto, H, Sasaki, S & Kumazawa, M *J geophys Res* **89,** 7852-7860 (1984)
17 Wood, B J & Kleppa, O J *Geochim cosmochim Acta* **45,** 529-534 (1981)
18 Akimoto, S & Fujisawa, H *J geophys Res* **73,** 1467-1479 (1968)
19 Ringwood, A E & Major, A *Phys Earth planet Inter* **3,** 89-108 (1970)
20 Yagi, T, Bell, P M & Mao, H-K *Yb Carnegie Inst Wash* **78,** 614-618 (1979)
21 Bina, C R & Wood, B J *J geophys Res* (submitted)
22 Anderson, O L, Schreiber, E, Liebermann, R C & Soga, N *Rev Geophys Space Phys* **6,** 491-524 (1968)
23 Birch, F *J geophys Res* **57,** 227-286 (1952)
24 Wood, B J & Holloway, J R *Geochim cosmochim Acta* **48,** 159-176 (1984)
25 Watt, J P, Davies, G F & O Connell, R J *Rev Geophys Space Phys* **14,** 541-563 (1976)
26 Sumino, Y & Anderson, O L in *CRC Handbook of Physical Properties of Rocks* Vol III (ed Carmichael, R S) 39-138 (CRC, Boca Raton, 1984)
27 Lees, A C, Bukowinski, M S T & Jeanloz, R *J geophys Res* **88,** 8145-8159 (1983)
28 Muirhead, K *J geophys Res* **90,** 2057-2059 (1985)
29 Ingate, S F, Ha, J & Muirhead, K J *Geophys J R astr Soc* **86,** 57-61 (1986)
30 Leven, J H *Phys Earth planet Inter* **38,** 9-27 (1985)
31 Silver, P G, Carlson, R W, Bell, P & Olsen, P *Eos* **66,** 1193-1198 (1985)

Natural and anthropogenic causes of lake acidification in Nova Scotia

Eville Gorham*, John K. Underwood†, Frank B. Martin‡ & J. Gordon Ogden III§

* Department of Ecology and Behavioral Biology,
University of Minnesota, Minneapolis, Minnesota 55455, USA
† Nova Scotia Department of Environment, Halifax,
Nova Scotia B3J 3B7, Canada
‡ Statistical Center, University of Minnesota, Saint Paul,
Minnesota 55108, USA
§ Department of Biology, Dalhousie University, Halifax,
Nova Scotia B3H 4J1, Canada

Controversy has arisen over the recent acidification of lakes[1–4], ascribed by many to anthropogenic acid deposition from the atmosphere[5,6], and by some to natural processes of soil acidification enhanced by the regrowth of forests after cutting and burning[7–10]. Here we show, by analysing the chemistry of Nova Scotian lakes and ponds on base-poor terrains, that both anthropogenic and natural acidification can be important We calculated correlations and regressions between hydrogen ion (H^+) concentrations and each of four predictors: dissolved organic carbon (DOC) (a surrogate for complex coloured organic acids, often of high molecular weight[11,12]), non-marine sulphate (denoted by the prefixed asterisk as $^*SO_4^{2-}$; a surrogate for acid deposition[13]), non-marine calcium ($^*Ca^{2+}$; the major basic cation from soil weathering and ion exchange), and the sum of the other non-marine base cations sodium, magnesium and potassium ($^*(Na^+ + Mg^{2+} + K^+)$). The results indicate that acidity in these waters is affected both by organic acids from peatland catchments and by acid deposition from long-range and local sources (see ref. 14).

Previous investigators ascribed acidity of lakes on diverse geological substrates in Halifax County to either peaty inflows[15] or acid deposition[16] Both used methods inadequate for measuring SO_4^{2-} in coloured waters[17], so another study was needed We sampled lakes on granite and quartzite terrains between 2 5 and 42 km around the Halifax-Dartmouth power plant, whose annual emissions declined from 21 5 kton SO_2 in 1980 to 8 3 kton in 1984 (ref 18) Thirty-seven surface waters were collected during 7-18 December 1984 after a dry period (September-November precipitation, 42% normal) and filtered through Reeve-Angel no 934AH glass-fibre filters before analysis H^+ was measured by a Radiometer combination electrode, DOC by photo-oxidation, gas dialysis and colorimetric titration with phenolphthalein, and Ca^{2+}, Mg^{2+}, Na^+ and K^+ by atomic absorption spectrophotometry SO_4^{2-} was analysed by ion chromatography and colorimetrically by methyl thymol blue after photo-oxidation, results were averaged (Ion chromatography yielded a mean 1 2% above that of colorimetry, r^2 was 0 97, the slope of the relation departed by only 2 5% from the 1 1 line, and the average difference between techniques was 6 μequiv $^{-1}$) $^*SO_4^{2-}$, $^*Ca^{2+}$ and $^*(Na^+ + Mg^{2+} + K^+)$ were estimated as differences between total and marine concentrations, the latter from ratios to chloride (Cl^-) in sea water Cl^- was measured by ion chromatography and colorimetrically by mercuric thiocyanate, results were averaged (Colorimetry yielded a mean 1 4% above that of ion chromatography, r^2 was 0 99, but the slope of the relation departed from the 1 1 line by 8%, so that at very low concentrations colorimetric values were 15% below those of ion chromatography, whereas at very high concentrations colorimetric values exceeded those of ion chromatography by almost 5% Average difference between techniques was 17 μequiv $^{-1}$) Colour was measured as absorbance at 320 nm in a 1-cm cell (A_{320}^{1cm}), and in Pt/Co units Because of plant uptake and microbial reduction, nitrate and ammonium ions—important components of acid deposition—were negligible (N usually <0 01 mg l^{-1}) Alkalinity was usually <2 and always <25 μequiv l^{-1} Ranges of chemical properties were DOC, 0 7-27 mg l^{-1}, colour (A_{320}^{1cm}), 0 001-0 574 and (Pt/Co units), 7-200, H^+, 1-98 μequiv l^{-1}, Cl^-, 90-832 μequiv l^{-1}, $^*SO_4^{2-}$, 12-203 μequiv l^{-1}, $^*Ca^{2+}$, 17-239 μequiv l^{-1}, and $^*(Na^- + Mg^{2+} + K^+)$, 2 4-105 μequiv l^{-1} In the last case, mean percentages were $^*Na^+$, 64, $^*Mg^{2+}$, 23, $^*K^+$, 13

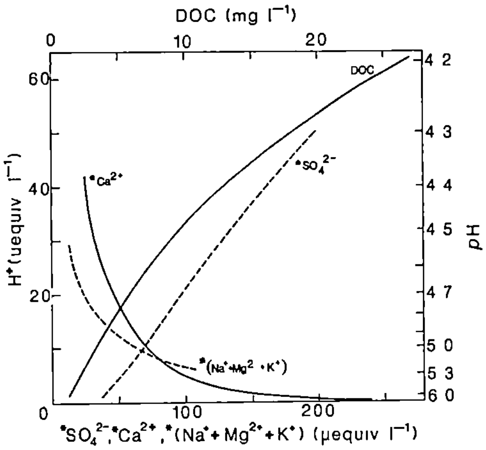


Fig 1 Partial regressions of H^+ on DOC, $^*SO_4^{2-}$, $^*Ca^{2+}$ and $^*(Na^+ + Mg^{2+} + K^+)$

We analysed the chemistry of 33 waters statistically, excluding four lakes by standard diagnostic criteria for regressions[19] Data were transformed to linearize relationships between H^+ and the predictors, error distributions were also normalized H^+ was transformed to its square root, other data to natural logarithms Transformed means were H^+, 3 93, DOC, 1 50, $^*SO_4^{2-}$, 4 43, $^*Ca^{2+}$, 4 03, and $^*(Na^+ + Mg^{2+} + K^+)$, 3 60

In the transformed scales, DOC is the best of the four predictors of H^+ its correlation coefficient is 0 746 and its partial correlation coefficient—in an equation including $^*SO_4^{2-}$, $^*Ca^{2+}$ and $^*(Na^+ + Mg^{2+} + K^+)$—is 0 908 (Except where specified, all correlation coefficients are significant at $p < 0 01$) For $^*SO_4^{2-}$ the corresponding statistics are lower, with a correlation

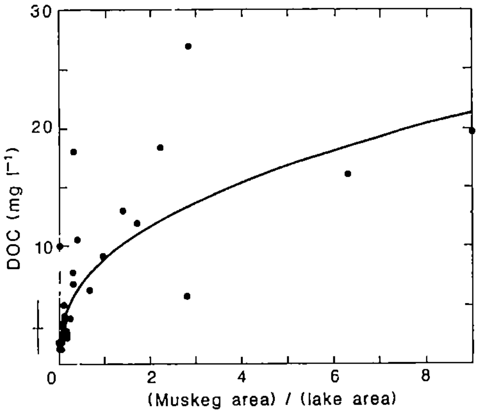


Fig. 2 Relation between DOC and the ratio (muskeg area)/(lake area) ($n = 30$). Cross at lower left indicates mean and range of sites without muskeg. Logarithmic regression, $\ln DOC = 0.395 \ln$ {muskeg area)/(lake area)} $+ 2.19$, correlation, $r^2_{ln} = 0.605$.

coefficient of 0.095 (not significant) and a partial correlation coefficient of 0.742. $^*Ca^{2+}$ predicts H^+ better than $^*SO_4^{2-}$, with correlation coefficients of -0.385 ($p < 0.05$) and -0.849 respectively, negative because of alkalinity associated with $^*Ca^{2+}$ released from soils and sediments. Coefficients for $^*(Na^+ + Mg^{2+} + K^+)$ shift sign after partial correlation, from 0.386 ($p < 0.05$) to -0.558. The multiple regression equation, including both natural and anthropogenic acidifying agents and neutralizing base cations as predictors, is

$$\sqrt{H^+} = 1.33 + 2.26 \ln DOC + 3.66 \ln {}^*SO_4^{2-} - 2.97 \ln {}^*Ca^{2+} - 1.40 \ln {}^*(Na^+ + Mg^{2+} + K^+) \qquad (1)$$

(DOC in mg l^{-1}, other variables in μequiv l^{-1}). The multiple correlation coefficient ($R^2 = 0.878$) accounts for most of the variance in H^+, and is distinctly greater than that ($R^2 = 0.823$) for H^+ on DOC, $^*SO_4^{2-}$ and $^*Ca^{2+}$.

Another possible cause of lake acidification is the deposition of sea salts, whose metal cations can displace adsorbed H^+ episodically and seasonally from soils and peats into streams and lakes[20]. Cl^- can be used as a predictor of the sea-salt effect on H^+, and when substituted for $^*(Na^+ + Mg^{2+} + K^+)$ in equation (1), it too yielded a correlation ($R^2 = 0.857$) distinctly greater than that with three predictors. However, when Cl^- was added to equation (1) as a fifth predictor, the correlation ($R^2 = 0.885$) did not improve appreciably.

Figure 1 illustrates partial relations of H^+ to DOC, $^*SO_4^{2-}$, $^*Ca^{2+}$ and $^*(Na^+ + Mg^{2+} + K^+)$ over their observed ranges of concentration, holding the other three predictors at their mean values in equation (1). DOC has a broader range of effect than $^*SO_4^{2-}$ on H^+, which rises from 1 to 64 μequiv l^{-1} over the full range of DOC increase and from 1 to 50 μequiv l^{-1} over the full range of $^*SO_4^{2-}$ increase. $^*Ca^{2+}$ has a negative effect upon H^+, which declines from 42 to <1 μequiv l^{-1} over the full range of $^*Ca^{2+}$ increase. The decline is curvilinear, reflecting the rapid drop in pH as bicarbonate alkalinity (associated primarily with $^*Ca^{2+}$) is exhausted. A smaller decrease in H^+, from 29 to 6 μequiv l^{-1}, occurs with increasing $^*(Na^1 + Mg^{2+} + K^+)$.

Correlations among DOC, colour and absorbance are highly significant, indicating humic substances, chiefly fulvic acids[11,12], produced by decomposition of plant remains in the catchments. DOC also correlates with the ratio (catchment area)/(lake area), because larger catchments generally contain larger areas of the peatlands that are the major sources of DOC input (see below) and the area of the lake is a major determinant of the volume in which the input is dissolved. The correlation between DOC and the ratio is logarithmic and highly significant ($r^2_{ln} = 0.392$). The prediction equation is $\ln DOC = (-0.095) + 0.634 \ln$ {(catchment area)/(lake area)}. Catchments have been mapped (by the Nova Scotia Department of Lands and Forests) as muskeg, rock

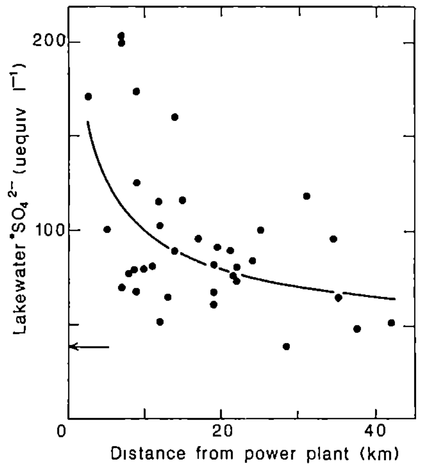


Fig. 3 Decline of lakewater $^*SO_4^{2-}$ outward from the Halifax-Dartmouth power plant ($n = 36$, rejecting one outlier). $\ln {}^*SO_4^{2-} = 5.35 - 0.321 \ln$ km, $r^2_{ln} = 0.266$. Horizontal arrow, background from long-distance transport.

barren, softwood (>75%) forest, hardwood (>75%) forest and mixed forest. Muskegs (which we remapped from air-photos, adjusting other areas appropriately) represent open peatland, sometimes with stunted trees, and are commonly dominated by *Sphagnum* moss. Rock barrens include large patches of ericaceous shrubs over acid mor humus on rock or shallow sandy soil, and small patches of *Sphagnum*. Muskegs and softwood forests on peat are the most likely sources of DOC to lakes[21,22]. Rock barrens are also possible sources because they often have patches of peat, and also lack deeper, less acid B soil horizons capable of precipitating humic substances with oxides of iron and aluminium[23].

For the 16 waters with all three cover types, analysis indicated significant correlation of DOC only with the ratio (muskeg area)/(lake area). For all 30 catchments with muskeg (Fig. 2), the logarithmic correlation and regression of these two variables are highly significant, the correlation accounting for 60% of the variance in DOC. Restriction to lakeside ($n = 14$) and lakeside plus streamside muskeg ($n = 26$) does not improve the correlation. The arithmetic scatter is, however, substantial, and softwood (black spruce and tamarack) swamps, which could not be mapped separately from upland softwood forests, deserve investigation in this context. The rate of water-flow through peatlands also influences DOC concentrations[21].

The $^*SO_4^{2-}$ in these lake waters comes chiefly from acid deposition. (Maps of soil and bedrock geology were inspected carefully to exclude the Halifax slate formation, in which local occurrences of pyrite are a source of $^*SO_4^{2-}$. However, there is a remote possibility that some $^*SO_4^{2-}$ is supplied through the use of crushed Halifax slate on gravel roads beyond the extent of its geological exposure[24].) Figure 3 shows the influence of Halifax-Dartmouth (population 191,000), an oil-fired generating station at Tufts Cove (360 MW) is the reference point for distance. The correlation between $^*SO_4^{2-}$ and distance is relatively low. Differences in vegetation cover, fluctuations in precipitation, wind speed and direction, and differences in water retention time and reduction of SO_4^{2-} in anoxic bottom waters also influence $^*SO_4^{2-}$ inputs and concentrations. The curve in Fig. 3 approaches a baseline (39μequiv $^{-1}$) representing long-distance transport of air pollutants, it was calculated by interpolating Halifax along a line connecting average $^*SO_4^{2-}$ concentrations in lakes at Kejimkujik National Park (46 μequiv l^{-1}) and Cape Breton Highlands National Park (20 μequiv l^{-1}), respectively 140 km SW and 350 km NE of Halifax[25].

Interaction of acid deposition with soil organic matter is claimed to decrease the solubility of humic compounds[8]. Halifax County, with a broad range of lake DOC and moderate levels of acid deposition, shows no evidence of this, the correlation between DOC and $*\mathrm{SO}_4^{2-}$ ($r_{\mathrm{ln}}^2 = 0\ 03$) being insignificant (see ref 26)

Coloured organic acids from catchments strongly influence the acidity of lakes on base-poor substrates, so that even small amounts of humic DOC affect lakewater pH greatly, between minimum DOC at $<1\ \mathrm{mg\,l^{-1}}$ and the log-normal mean at $4\ 5\ \mathrm{mg\,l^{-1}}$, pH declines more than one unit to 4 8 (Fig 1) This conclusion probably applies throughout Nova Scotia, 769 lakes, mostly on base-poor substrates, exhibit a log-normal distribution of colour (J K U unpublished observations) greater than that of Halifax County lakes (antilog mean = 35 Pt/Co units in Nova Scotia, 27 units in Halifax County) Moreover, in 234 Nova Scotian lakes[27], organic anions average 49 μequiv $\mathrm{l^{-1}}$ and $*\mathrm{SO}_4^{2-}$ 60 μequiv $\mathrm{l^{-1}}$ The influence of coloured organic acids is likely in other areas (for example much of the Precambrian Canadian Shield) on comparably base-poor substrates with an abundance of acid peatlands, and with similar concentrations of organic anions (see ref 28) Their significance for lake-acidification models[6,29] is not known and should be evaluated

Despite the influence of coloured organic acids from natural sources (see refs 14, 26 and 30), the effect of acid deposition on surface waters in Nova Scotia is clear (Fig 3) and especially important at snow-melt and other periods of high flow[14,26] It has probably caused naturally acidified waters to acidify further and beyond thresholds for losses of diatoms[31] and salmon[32,33] The degree of natural acidification, as well as the amount of wet $*\mathrm{SO}_4^{2-}$ deposition from the atmosphere, should be taken into account in determining target loadings for acid deposition

We thank the National Science Foundation and the Nova Scotia Department of the Environment for support, R F McCurdy, M V Santelmann, P Y Schwartz and D L Smith for assistance, and S E Bayley, P J Dillon, F C Elder, J N Galloway, L Hedin, H F Hemond, G E Likens, S A Norton and D W Schindler for advice

Received 16 June, accepted 30 September 1986

1 Johnson, N M, Likens, G E, Feller, M C & Driscoll, C T *Science* **225**, 1424-1425 (1984)
2 Seip, H & Dillon, P J *Science* **225**, 1425-1426 (1984)
3 Wright, R F *Science* **225**, 1426, 1428 (1984)
4 Krug, E C & Frink, C R. *Science* **225**, 1428, 1433-1423 (1984)
5 Drablos, D & Tollan, A (eds) *Ecological Impact of Acid Precipitation* (SNSF Project, Oslo, 1980)
6 Linthurst, R A (ed) *The Acidic Deposition Phenomenon and its Effects Critical Assessment Review Papers*, vol 2, *Effects Sciences* (US EPA 60018 83 016AF, 1984)
7 Rosenqvist, I T *Sci tot Envir* **10**, 39-40 (1978)
8 Krug, E C & Frink, C R. *Science* **221**, 520-525 (1983)
9 Lefohn, A S & Klock, G E *J Air Pollut Control Ass* **35**, 632-637 (1985)
10 Krug, E C, Isaacson, P J & Frink, C R *J Air Pollut Control Ass* **35**, 109-114 (1985)
11 Oliver, B G, Thurman, E M & Malcolm, R L *Geochim cosmochim Acta* **47**, 2031-2035 (1983)
12 McKnight, D, Thurman, E M, Wershaw, R L & Hemond, H *Ecology* **66**, 1339-1352 (1985)
13 Gorham, E, Martin, F B & Litzau, J T *Science* **225**, 407-409 (1984)
14 Kerekes, J in *Report, Fall 1984 Workshop, Atlantic Region LRTAP Monitoring and Effects Working Group* (ed Taylor, B L) 84-85 (Atmos Environ Service, Bedford, Nova Scotia, 1984)
15 Gorham, E *Limnol Oceanogr* **2**, 12-21 (1957)
16 Watt, W D, Scott, D & Ray, S *Limnol Oceanogr* **24**, 1154-1161 (1979)
17 Gorham, E & Detenbeck, N E *J Ecol* **74**, 899-903 (1986)
18 Underwood, J K Ph D thesis (Tech Univ Nova Scotia, 1984)
19 Weisberg, S *Applied Linear Regression* 2nd edn (Wiley, New York, 1985)
20 Christopherson, N, Seip, H M, Qvenild, C & Tollan, O *Nordic Hydrol.* **13**, 105-114 (1982)
21 Glaser, P H, Wheeler, G A, Gorham, E & Wright, H E *J Ecol.* **69**, 575-599 (1981)
22 Gorham, E, Eisenreich, S J, Ford, J & Santelmann, M V in *Chemical Processes in Lakes* (ed Stumm, W) 339-363 (Wiley, New York, 1985)
23 McDowell, W H & Wood, T *Soil Sci.* **137**, 23-32 (1984)
24 King, M *Acid Drainage and the Acidification of Nova Scotia Waters* (preprint, Water Planning and Management Branch, Inland Waters Directorate, Environment Canada, 1985)
25 Underwood, J K, Ogden III, J G, Kerekes, J & Vaughan, H H *Wat Air Soil Pollut* (in the press)
26 LaZerte, B D & Dillon, P J *Can J Fish aquat Sci* **41**, 1664-1677 (1984)
27 Underwood, J K, Ogden III, J G & Smith, D L *Wat Air Soil Pollut* (in the press)
28 Jeffries, D S, Wales, D L, Kelso, J R M & Linthurst, R A *Wat Air Soil Pollut* (in the press)
29 Wright, R F in *Meteorological Aspects of Acid Rain* (ed Bhumralkar, C M) 213-230 (Butterworth, Boston, 1984)
30 Jones, M K, Marmorek, D R., Reuber, B S, McNamee, P J & Rattie, L P *Brown Waters Relative Importance of External and Internal Sources of Acidification on Catchment Biota, Review of Existing Knowledge* (ESSA Ltd, Toronto, Ontario, 1986)
31 Vaughan, H H, Underwood, J K & Ogden III, J G *Wat Air Soil Pollut* **18**, 353-361 (1982)
32 Farmer, G J, Goff, T R, Ashfield, D & Samant, H S in *Acid Rain and the Atlantic Salmon* (ed Sochasky, L) 73-91 (Int Atlantic Salmon Found, Special Publ no 10, 1981)
33 Watt, W D *Wat Air Soil Pollut* (in the press)

A 40,000 year-old human occupation site at Huon Peninsula, Papua New Guinea

Les Groube*, John Chappell†, John Muke*
& David Price

* Department of Anthropology and Sociology,
University of Papua New Guinea, Port Moresby, New Guinea
† Biogeography and Geomorphology, Research School of Pacific Studies, Australian National University, Canberra, Australia
‡ Department of Physics, The Faculties,
Australian National University, Canberra, Australia

The geographical position of the island of New Guinea suggests that it may have been an early staging post in the Pleistocene settlement of Australia from the Indonesia–Indochina region. Previous data have not supported this, as archaeological sites 35,000 to 40,000 years old occur in southern Australia[1,2], whereas the earliest previously known in Papua New Guinea is 26,000 years old[3–5]. We now report evidence that the north coast of Papua New Guinea was occupied at least 40,000 years ago Sahuland, which is the greater land area of Australia and New Guinea plus their connecting continental shelf exposed as land when Pleistocene sea levels were lower than now[6,7], was occupied by humans in several widely separated areas at that time. A distinctive 'waisted axe' culture appears to have existed in New Guinea and probably in Australia in the Late Pleistocene, but antecedents are not yet known from east and southeast Asia. There is evidence for hafting of these tools at a date which is earlier than known elsewhere in the world.

Our evidence is found at the southeast end of Huon Peninsula (Fig 1), where a flight of raised coral terraces preserves the former coastal environment of a succession of Upper Pleistocene stages The terraces have been described and dated by ^{14}C and ^{230}Th/^{234}U methods, showing continual tectonic uplift of the coast[8–11] Each coral reef terrace formed when a glacio-eustatic rise of sea level overtook the rising land[9] During each period of reef formation, the coastal environment included local lagoons and fringing reefs By analogy with the present coastal fringe with its Holocene reefs, these provided a productive and hospitable habitat for human occupation[9,12] Figure 1 shows the terrace flight in the vicinity of the archaeological site described here

The archaeological site occurs in a stream gully running into a small palaeo-lagoon on the reef complex labelled IIIa by Chappell[9], which is dated as 45,000 to 53,000 years by ^{230}Th/^{234}U This age and the ^{230}Th/^{234}U chronology of the entire terrace sequence is supported by correlation with deep-sea core records[13] In the area shown in Fig 1, reef IIIa overlies seaward-dipping bedded sandy tuffs that are exposed in several streams and are seen to pass inland beneath the front of the next reef higher in the sequence A set of three strongly weathered tephras occur in the depression behind reef IIIa and on the lower front of the rise to reef IVb (Fig 1) Artefacts found in the excavation described below occur within the tephra sequence The tephras are considerably altered by pedogenesis All are clay-like and the upper two are dark brown to black, and each contains fine sand to fine silt-size angular tephric quartz in greater concentrations towards their bases The tephra pile varies from 1 5-2 8 m thick due to slope wash and redeposition after each fall, as well as to downslope creep which is indicated by small slip planes visible in the excavation

The age of the tephras is critical to our conclusions about the age of the artefacts The tephras do not occur on 6,000-year-old (Holocene) reef I, which is broad and contains potential tephra-trapping palaeo-lagoons (Fig 1) They have not been found on reef II, which is dated to 28,000 to 30,000 years[9,11], nor on reef IIIb which is 40,000 years old[9–11], although the possibility they

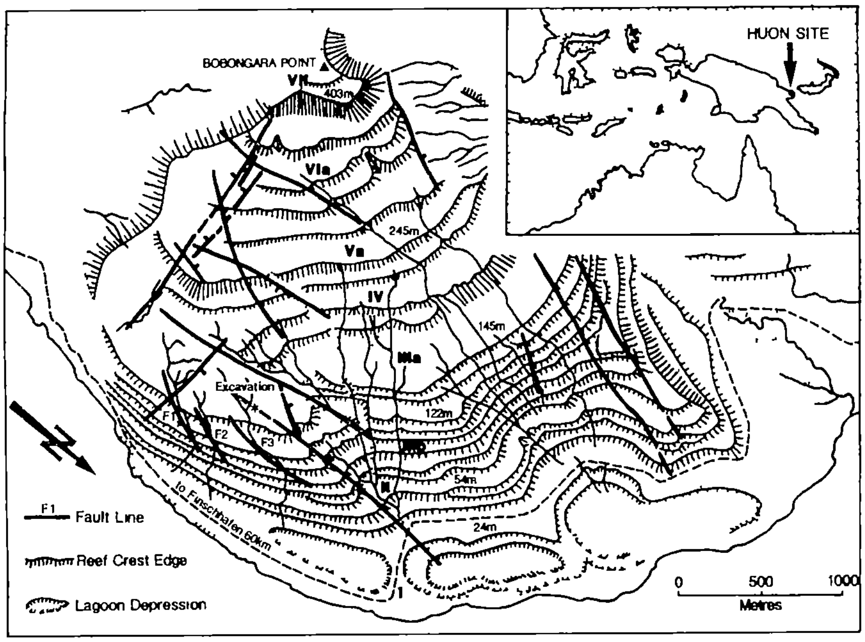


Fig 1 Bobongara area of Huon Peninsula coral terraces Archaeologic site marked * Excavation Inset shows land exposed at times of Pleistocene low sea levels (white areas), and location of Huon Peninsula site in Papua New Guinea

were on these reefs and subsequently were stripped cannot be eliminated The inference that the tephras are older than 40,000 years is supported by thermoluminescence (TL) dating of the quartz particles within each unit This dating was done at Australian National University, using the fine-grain technique[14] All samples show well-defined TL plateaux from about 350° to 475 °C Laboratory-induced TL growth curves were linear, and fading tests (1 month) found no TL loss Hence, we are confident of the calculated archaeologic doses Dose rate, based on laboratory analyses of potassium (K) and total α count rate for each tephra, is less certain due to a peculiarity in the K-values Measured K_2O content of the tephras ranges from 0 05–0 11% This is very low for tephra, and K-leaching is suspected to have occurred during strong pedogenesis The nearby Bismark volcanics average K_2O of 1 2%[15] Hence, dose rates are estimated on the basis of (1) measured K_2O, and (2) $K_2O = 1\ 2\%$ The latter is almost certainly greater than the average value over the history of the deposits The alpha dose rate is taken as measured, which may also be higher than the historical average due to thorium growth in the neighbouring reefs, which are in U-series disequilibrium[8,10]

TL data and calculated ages are shown in Table 1 Ages based on measured K_2O and assumed K_2O (=1 204%) are both consistent with stratigraphic order, showing that T1 is significantly younger than T2 and T3, which are similar to each other Ages based on measured K_2O are higher than those based on assumed K_2O because the K_2O values are lower We believe that ages based on assumed K_2O are almost certainly too low, first because they are based on much higher K_2O than measured, and second because the assumption that $H_2O = 0$ is unrealistic as annual rainfall at the site is 2,500 mm and the soil has been found to be generally wet during excavation Soil moisture reduces dose rate, and the zero H_2O assumption leads to unduly low age estimates We conclude that the T2 and T3 are at least 40,000 years old, and probably T1 as well, consistent with their occurrence on Reef IIIa but not on the younger reef terraces

Stratigraphy of the archaeological excavation is shown in Fig 2 The site is adjacent to an ephemeral channel (Jo's Creek) which descends the front of reef IVb and crosses the depression behind the crest of reef IIIa Two sections were excavated, one along the southeast creek bank and the other at right angles to this ('trench section' in Fig 2) The creek section was excavated as a vertical face cut slowly back into the slope from the channel, and the trench was cut slowly downwards maintaining a horizontal floor until basal tuff and limestone was encountered Sieving was not possible in the sticky clay-like tephras, and most work was by trowel Figure 2 shows that the tephras are continuous but vary in thickness Some variation is due to mass creep, as mentioned, and some reflects infilling of temporary channels which are thought to have formed by runoff before and between tephra falls Such a channel is seen beneath findspot 2 in the trench section (Fig 2)

Artefacts were recovered at several points indicated in Fig 2, the most important being at the interface between T2 and T3 at findspots 1 and 2 These and that from findspot 3 are waisted axes, described shortly A stone core and 2 flakes were found at and near findspot 4 All were recovered from *in situ* positions by direct excavation The waisted axe at findspot 2 was in several pieces with interstitial tephra T2, attesting to burial after breakage with little subsequent displacement The waisted axe from

Table 1 Thermoluminescence dating results

			TL age (yr BP)			
Tephra unit	Archaeological dose (rad)	Measured $K_2O\%$	Minimum ADR (rad yr^{-1})	Age*	Maximum ADR (rad yr^{-1})	Age
T1	12,600	0 0543	0 299	42,100	0 402	31,300
T2	9,430	0 1076	0 156	60,500	0 254	37,100
T3	8,732	0 1372	0 146	59,800	0 241	36,200

Column 5 shows ages based on measured K_2O and column 7 ages based on assumed $K_2O = 1\ 204\%$ Other assumed parameters are Rb = 110 ppm, $H_2O = 0$, and cosmic dose = 0 015 rads yr^{-1} Standard errors of measurement are not shown as these are much less than K_2O uncertainty Tephra units are labelled T1 to T3 in descending stratigraphic order

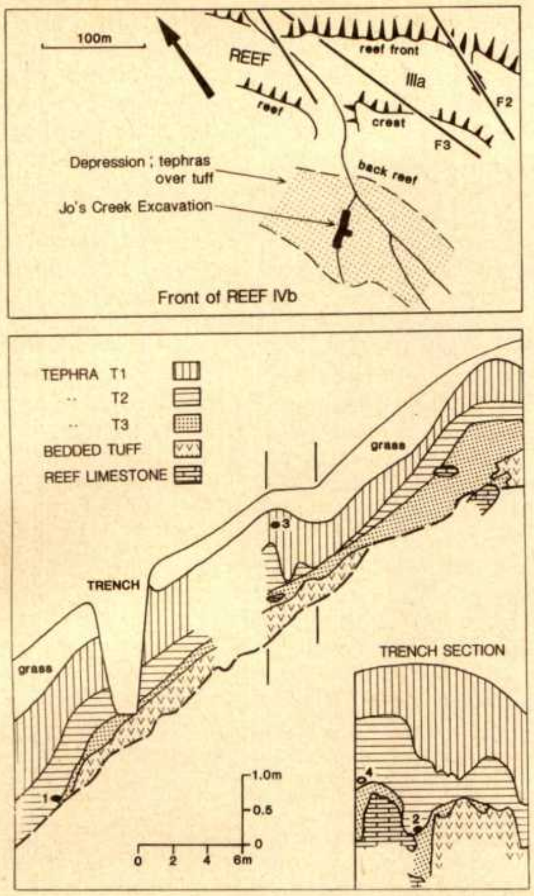


Fig. 2 Location of excavation (Jo's Creek) on Terrace IIIa at Bobongara, and stratigraphic sections at excavation. Main section follows creek; trench section at right angles to this. Numbered spots show locations of artefacts described in text.

spot 1, shown in Fig. 3*a*, is heavy (1.68 kg) and is marked by a distinct groove between the waisting notches. It is made of 2-pyroxene andesite comparable with boulders in the nearby Masaweng River. It has a skin of semi-opalline cristobalite (X-ray diffraction determination) and a thin section from a 1-cm core from the underside shows the albitic groundmass to be altered to microcrystalline quartz. These features are consistent with long burial in a weathering tephric environment. We refer to the tool as a waisted 'axe' on the grounds of its similarity in terms of size, weight, strength, and edge characteristics to axes. We interpret the groove across the waist as due to haft-lashing, by analogy with ethnographic examples from Melanesia[16] and Australia.

In addition to the artefacts found *in situ*, over 100 waisted axes of similar dimensions and pattern, although without hafting grooves, were found loose in creek beds around the site, mostly on or somewhat above reef IIIa. All are unifacial, of similar material to that in Fig. 3*a*, apparently manufactured by spallation from rounded river boulders followed by flake retouching, often very steep, of the flake-scar surface. A typical example is

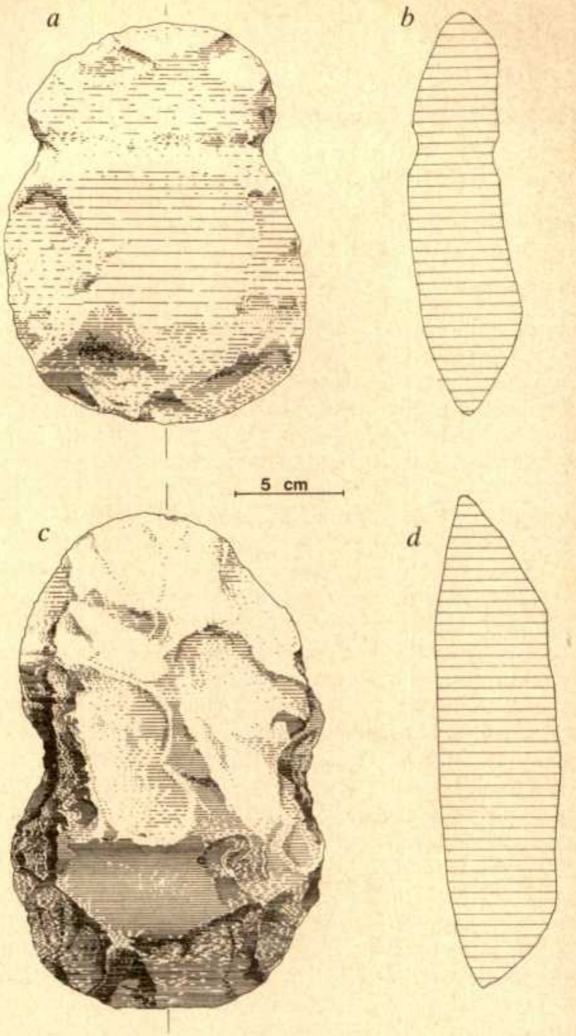


Fig. 3 *a, b*, Grooved waisted axe from findspot 1 in excavation. *c, d*, Typical non-grooved waisted axe from near to excavation.

shown in Fig. 3*b*. Other loose artefacts in creeks at or above the site include a chert core, various flakes, and one 'stemmed' and one 'knobbed' axe, both unifacial and made of andesite. Our waisted axes have similarities with artefacts discovered elsewhere. In highland New Guinea, waisted axes at 26,000 years are reported at Kosipe[4] and occur from 6,000 to over 10,000 BP at Yuku[17]. Pleistocene ages are inferred for waisted axes at Kangaroo Island in south Australia and at Mackay in north Queensland[18]. Waisted axes of inferred Holocene age occur at Botal Tobago[19] and Japan[20], in east Asia, although these appear to be smaller than the Huon Peninsula tools.

We thank the people of Hubegong village and the Tewai local government for assistance, the University of Papua New Guinea for support and the Australian Department of Foreign Affairs for financial assistance. We also are grateful to Professor J. Golson for assistance with the manuscript.

Received 14 May; accepted 30 September 1986.

1. Bowler, J. M. in *The Origin of the Australians* (eds Kirk, R. L. & Thorne, A. C.) 55–77 (Australian Institute of Aboriginal Studies, Canberra, 1976).
2. Pearce, R. H. & Barbetti, M. *Archaeol. Oceania* **16**, 173–178 (1981).
3. Hope, J. S., Golson, J. & Allen, J. *J. hum. Evol.* **12**, 37–60 (1983).
4. White, J. P., Crook, K. A. W. & Ruxton, B. P. *Proc. prehist. Soc.* **36**, 152–170 (1970).
5. Gillieson, D. S. & Mountain, M. J. *Archaeol. Oceania* **18**, 53–62 (1983).
6. Chappell, J. in *The Origin of the Australians* (eds Kirk, R. L. & Thorne, A. G.) 11–22 (Australian Institute of Aboriginal Studies, Canberra, 1976).
7. Birdsell, J. B. in *Sunda and Sahul* (eds Allen, J., Golson, J. & Jones, R.) 113–167 (Academic, London, 1977).
8. Veeh, H. H. & Chappell, J. *Science* **167**, 862–865 (1970).
9. Chappell, J. *Bull. Geol. Soc. Am.* **85**, 553–570 (1974).
10. Bloom, A. L., Broecker, W. S., Chappell, J., Matthews, R. K. & Mesolella, K. J. *Quat. Res.* **4**, 185–205 (1974).
11. Chappell, J. *Search* **14**, 99–101 (1983).
12. Chappell, J. *Archaeol. Oceania* **17**, 69–78 (1982).
13. Chappell, J. & Shackleton, N. J. *Nature* (in the press).
14. Fleming, S. *Thermoluminescence Techniques in Archaeology*, 58–80 (Clarendon, Oxford, 1979).
15. Johnson, R. W., Mackenzie, D. E., Smith, I. E. & Taylor, G. A. M. in *The Western Pacific* (ed. Coleman, P. J.) 523–533 (University of Western Australia Press, 1973).
16. Parkinson, R. H. R. *Dreissig Jahre in der Sudsee*, 405–407 (Strecker & Schroder, Stuttgart, 1907).
17. Bulmer, S. *Oceania* **34**, 246–265 (1964).
18. Lampert, R. J. *Rec. Austr. Mus.* (*Sydney*) No. 135 (1983).
19. Leach, E. R. *Man*, **38**, 161–163 (1938).
20. Chard, C. S. *Northeast Asia in Prehistory* (University of Wisconsin Press, 1974).

Reproductive failure in common seals feeding on fish from polluted coastal waters

Peter J. H. Reijnders

Research Institute for Nature Management, Department of Estuarine Ecology, Postbus 59, 1790 AB Den Burg, Texel, The Netherlands

The population of common seal *Phoca vitulina* in the westernmost part of the Wadden Sea, The Netherlands, has collapsed during the past few decades[1,2]. Between 1950 and 1975 the population dropped from more than 3,000 to less than 500 animals. Comparative studies of common seal populations from different parts of the Wadden Sea reveal that pup production has declined sharply only in the western (Dutch) part[2-4]. A comparative toxicological study[5] on the levels of heavy metals and organochlorines in tissues of seals from the western and northern parts of the Wadden Sea shows that only the polychlorinated biphenyl (PCB) levels differ significantly. This is predominantly a result of PCB pollution from the river Rhine[6-9], which mainly affects the western (Dutch) part. PCBs are thought to be responsible for the low rate of reproduction in Dutch common seals on the basis of epidemiological and experimental data on the ability of PCBs to interfere with mammalian reproduction[5,10]. Here I report that reproductive failure in common seals from the Dutch Wadden Sea is related to feeding on fish from that polluted area. This is the first demonstration of a causal relationship between naturally occurring levels of pollutants and a physiological response in marine mammals.

An experiment with two groups of 12 female common seals was carried out to test the detrimental effects of PCBs on seal reproduction. Each group, consisting of 7 seals from the east coast of the United Kingdom and 5 seals from the Museum of Natural History at Texel, was fed a diet containing different levels of pollutants. Group 1 received fish (predominantly plaice, flounder and dab, with some eelpout and hooknose) caught in the western part of the Wadden Sea. Group 2 received fish (mainly mackerel) from the north-east Atlantic. The fish were maintained at −28° C and thawed before being fed to the seals. The diets are comparable with respect to nutritional quality, except for fat levels; this was compensated for by a high daily intake. Residue analysis for aldrin, dieldrin, endrin, heptachlor, hepox, α,β,γ-hexachlorocyclohexane, pentachlorobenzene, hexachlorobenzene, *pp'*-dichlorodiphenyl-dichloroethylene (DDE), *op'*-dichlorodiphenyl-dichloroethane, *pp'*-dichlorodiphenyl-dichloroethane and PCBs showed statistically significant differences between the two diets for PCBs and *pp'*-DDE. The average daily intake (during ~2 years) was 1.5 mg PCBs and 0.4 mg *pp'*-DDE for group 1, and 0.22 mg and 0.13 mg for group 2 (J. P. Boon, P.J.H.R., J. Dols, and P. F. Wensvoort, unpublished data). Three males receiving Atlantic fish were alternated between both groups during the mating period. To detect whether hormonal regulation was affected by pollutants, blood samples were taken regularly and serum concentrations of progesterone and oestradiol-17β were determined. Because the timing of oestrus differs for individuals, all profiles have been synchronized by taking the day of previous delivery as day zero and adjusting the profiles of non-pregnant animals accordingly. Although there is some published information about the hormonal changes during reproduction in seals[11,13], none of these studies reported the changes during a complete annual reproductive cycle in a specific group of females. The oestradiol-17β and progesterone concentration profile of the control group (Fig. 1*e,g*) illustrate the normal annual cycle. The number of non-pregnant animals in the control group is too small to consider their reproductive cycle as 'normal' for non-pregnant females.

No statistically significant differences (Spearman, $p < 0.005$) were found in the progesterone and oestradiol-17β profiles of pregnant seals from groups 1 and 2. Identical tests for the non-pregnant animals in both groups gave the same result. This is important as it implies that the differences in diet do not influence hormone patterns. The patterns resemble those for eutherian mammals in general[14,15]. Pinnipeds are classified as having an obligate and seasonal embryonic diapause[16,17]. In common seals in this study implantation probably occurred at

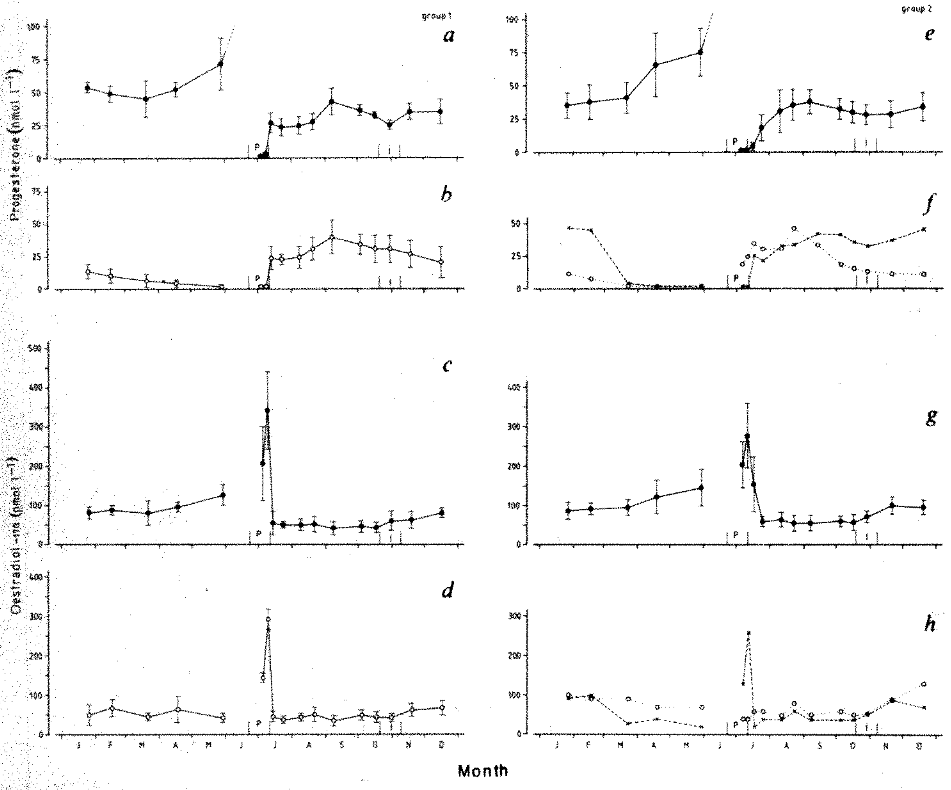


Fig. 1 Circulating serum concentrations of progesterone and oestradiol-17β during the reproductive season 1983–84, in common seals feeding on fish from the Wadden Sea (group 1) and on Atlantic fish (group 2). *a–d*, Group 1 (*a,b,c,d*): (● - - - ●), pregnant seals ($n = 4$); (○—○), non-pregnant seals ($n = 8$). *e–h*, Group 2: (● - - - ●), pregnant seals ($n = 10$); (× - - - ×) and (○—○) non-pregnant seals. Vertical bars, 1 s.d.; p = period over which previous parturitions occurred; i = implantation period. Blood samples were obtained from veins in the hind flippers after restraining the animals on a V-shaped bench. The blood (~8 ml) was collected in vacutainers, centrifuged and the serum stored frozen at −20 °C. Serum concentrations of progesterone and oestradiol-17β were determined by radioimmunoassay. The intra run coefficients of variation for progesterone and oestradiol-17β were consistently below 3.0% and 3.05% respectively.

the end of October or beginning of November, and was characterized by an increase in oestradiol-17β levels (Fig. 1, *c*, *g*).

The reproductive success was significantly lower in group 1 than in group 2 (Table 1; $p<0.02$, Fisher's exact probability test). The hormone profiles of the non-pregnant animals in group 1 show that the effect occurs at a stage of the reproductive process around implantation; the follicular, luteal and post-implantation phases until the end of gestation are not affected. These findings corroborate the results from experiments with mink[18,19] where PCBs impaired reproduction: ovulation, mating and implantation occurred but were followed by early abortion or resorption. These results may be due to hormonal disturbance, to direct dominant-lethal action or to an embryo lethal effect caused by toxicants. Hormonal disturbance may be caused by disruption of the steroid synthetic pathway resulting in reduced circulating levels of hormones. PCBs are known to cause microsomal enzyme induction, accelerating hydroxylation of body steroids such as oestrogens[5]. To determine if pollutants suppress circulating levels of hormones during the implantation period, the four subgroups (pregnant and non-pregnant seals in groups 1 and 2) are compared. A statistical test revealed no difference between progesterone levels of pregnant animals in group 2 and non-pregnant animals in group 1. It is concluded that progesterone levels were not reduced in group 1. The rise in oestradiol levels of the non-pregnant seals in group 2—indicating follical growth—is lacking in non-pregnant seals of group 1, which suggests that the nature of non-pregnancy differs between the groups. Because there were only 2 non-pregnant animals in group 2, this could not be tested statistically. The results also demonstrate that the levels of oestradiol in all group 1 seals are lower than those in group 2 combined, ($p<0.05$, Wilcoxon), although the initial rise in some of the animals of group 1 (Fig. 1*c*) was apparently sufficient to ensure reproductive success. The mechanisms behind the smaller increase in oestradiol levels and its consequences for the priming effect on the endometrium, as well as the maternal rejection response[20-22], will be discussed in detail elsewhere. Hypotheses about impaired steroid binding capacity by PCBs[23], a dominant-lethal action and an embryo lethal effect cannot be tested with the information available.

I conclude that the reproductive success of the seals receiving the diet with the highest level of pollutants was significantly decreased. No circumannual reduction in levels of circulating hormone levels was observed. The reproductive process is disrupted in the post-ovulation phase. The period around implantation seems to be the most sensitive stage, but no conclusions about the mechanism of action can be drawn. A similar experiment was simultaneously carried out with the American mink *Mustela vison*, a fish-eating mammal with a comparable reproductive physiology. It was designed to test whether pure PCBs had the same effect as the PCB-polluted fish. Three groups of 'Standard' type mink were fed either a basic diet of commercial mink cereal, or the same basic diet supplemented with livers from fish caught in the western part of the Wadden Sea or with Clopen A-60 or A-30. The results show that reproduction is inhibited at very low (25 μg per day) levels of PCB intake and that the effects of the pure PCB diet were identical to those of contaminated fish diet[18].

Irrespective of the precise mechanism involved, the results from this study show that the reproductive failure in common seals from the Dutch Wadden Sea is related to feeding on fish from that polluted area. The available epidemiological experimental data on effects and levels of PCBs in seals and mink fed on fish from this area[5,27] suggest that these organochlorines are the main cause of this failure.

I thank A. Grijzen, H. Jansens and G. van Weert of De Wever Hospital, Heerlen, for carrying out the radioimmunoassays; J. Schoonheden for the illustrations, M. C. Vleugel for care of the animals and J. Zegers for his assistance throughout the whole experiment.

1. Reijnders, P. J. H. *Neth. J. Sea Res.* **10,** 223–235 (1976).
2. Reijnders, P. J. H. *Neth. J. Sea Res.* **12,** 164–179 (1978).
3. Reijnders, P. J. H., Drescher, H. E., van Haaften, J. L., Bøgebjerg Hansen, E. & Tougaard, S. in *Marine Mammals of the Wadden Sea* (eds Reijnders, P. J. H. & Wolff, W. J.) 19–32 (Balkema, Rotterdam, 1981).
4. Reijnders, P. J. H. *Z. Säugetierk.* **48,** 50–54 (1983).
5. Reijnders, P. J. H. *Neth. J. Sea Res.* **14,** 30–65 (1980).
6. Duinker, J. C. & Hillebrand, M. T. J. *Neth. J. Sea Res.* **13,** 256–281 (1979).
7. Duinker, J. C., Hillebrand, M. T. J., Nolting, R. F. & Wellershaus, S. *Neth. J. Sea Res.* **15,** 141–169 (1982).
8. Duinker, J. C., Hillebrand, M. T. J., Nolting, R. F. & Wellershaus, S. *Neth. J. Sea Res.* **15,** 170–195 (1982).
9. Duinker, J. C., Boon, J. P. & Hillebrand, M. T. J. *Neth. Inst. Sea Res. Publ. Ser.* **10,** 211–228 (1984).
10. Reijnders, P. J. H. in *Marine Mammals of the Wadden Sea* (eds Reijnders, P. J. H. & Wolff, W. J.) 38–47 (Balkema, Rotterdam, 1981).
11. Daniel, J. C., Jr. *J. Fish. Res. Bd Can.* **33,** 65–66 (1975).
12. Boyd, I. L. *J. Repr. Fert.* **69,** 157–164 (1983).
13. Raeside, J. I. & Ronald, K. *J. Reprod. Fert.* **61,** 135–139 (1981).
14. Amorsoso, E. C., Heap, A. B. & Renfree, M. B. in *Hormones and Evolution* (ed. Barrington, E. J. W.) 925–989 (1979).
15. Flint, A. P. F. *J. Repr. Fert. Suppl.* **29,** 215–227 (1981).
16. Renfree, M. B. & Calaby, J. H. *J. Reprod. Fert. Suppl.* **29,** 1–9 (1981).
17. Boshier, D. P. *J. Repr. Fert. Suppl.* **29,** 143–149 (1981).
18. Boer, M. den in *A. Rep. RIN* (ed. Rossum, T. van) 77–86 (Rijksinstituut voor Natuurbeheer, Leersum, 1983).
19. Jensen, S., Kilström, J. E., Olsson, M., Lundberg, C. & Orberg, J. *Ambio* **6,** 239 (1977).
20. Finn, R., St. Hill, C. A., Govan, A. J. Ralfs, I. G. & Gurney, F. J. *Br. Med. J.* **3,** 150–152 (1972).
21. Anderson, R. H. & Munroe, C. W. *Am. J. Obstet. Gynec.* **84,** 1096–1103 (1962).
22. Grossman, C. J. *Science* **227,** 257–261 (1985).
23. Bitman, J. & Cecil, H. C. *J. Agric. Fd Chem.* **18,** 1108–1112 (1970).
24. Osborn, D. B., Cooke, A. S. & Freestone, S. *Envir. Pollut.* **25,** 305–319 (1981).
25. Ronald, K., Foster, M. E. & Johnson, E. *Can. J. Zool.* **47,** 461–468 (1969).
26. Bouckaert, P. X. J. M. thesis (Univ. Nijmegen, Netherlands, 1984).
27. Reijnders, P. J. H. *Mar. Mamm. Sci.* (submitted).

Table 1 of this paper appears on page 418.

A common mammalian plan of accessory optic system organization revealed in all primates

H. M. Cooper & M. Magnin

Laboratoire de Neuropsychologie Expérimentale, INSERM U-94,
16 Avenue Doyen Lépine, 69500 Bron, France

The accessory optic system (AOS), which was described as early as 1870 by Gudden[1], constitutes a distinct midbrain visual pathway in all classes of vertebrates[2]. In non-primate mammals, retinal fibres of this system project to a set of three nuclei[3,4]: the dorsal (DTN), the lateral (LTN) and the medial (MTN) terminal nuclei. Whereas all AOS cells respond to the slow motion of large visual stimuli, the neurons are tuned to complementary directions of movement[5]: horizontal temporo-nasal direction for the DTN, vertical up and down for the LTN and vertical down for the MTN. It has thus been suggested that these nuclei establish a system of retinal coordinates for the detection of whole field motion[6]. As the AOS provides direct and indirect pathways to both oculomotor and vestibular structures[7,8], each of these nuclei is thought to be an essential link in the co-ordination of eye and head movements in relation to movement within the visual field. One problem for the generalization of this theory is that the medial terminal nucleus has never been found in primates. In this report we establish both the existence of this nucleus and its afferent input from the retina in all major groups of primates (prosimians, New and Old World monkeys and apes), indicating a common anatomical plan of organization of the AOS in mammals.

We used both autoradiographic and histochemical anterograde tracing techniques to study the retinal projection to the AOS in the two primate suborders[9] Strepsirhini (prosimians: lemurs and lorises) and Haplorhini (simians: New and Old World monkeys, apes and man). For the Strepsirhines we studied five mouse lemurs (*Microcebus murinus*) and one bushbaby (*Galago demidovii*). Haplorhine species included two marmosets (*Callithrix jacchus*), three macaques (*Macaca fascicularis*) and two gibbons (*Hylobates concolor*). Most animals were injected with a radioactive amino acid mixture (500–

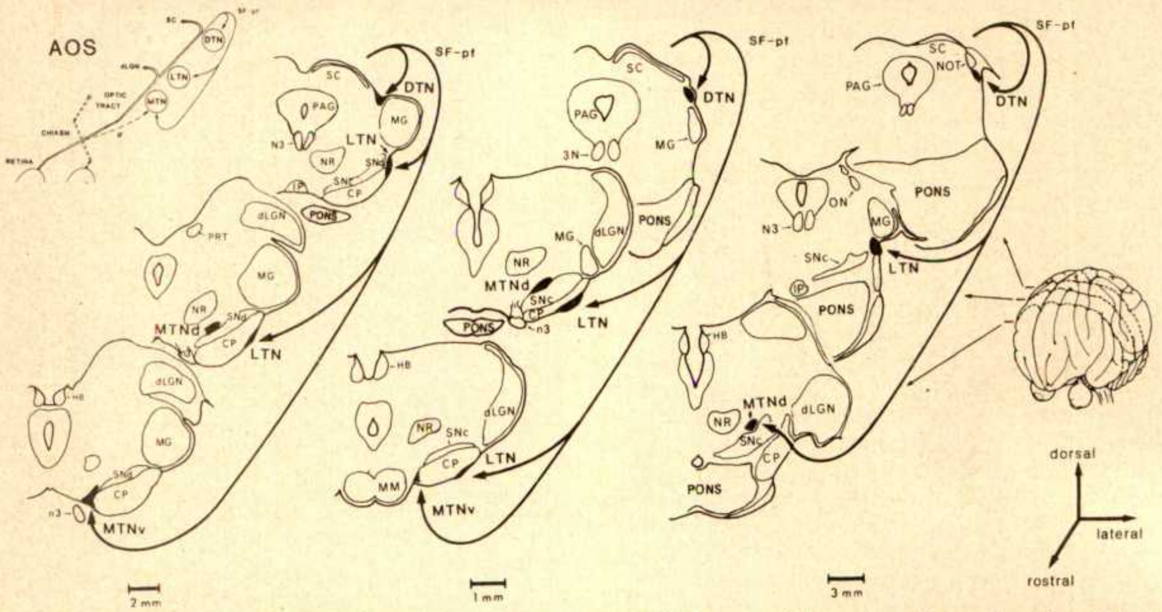


Fig. 1 Organization of the accessory optic system compared in a non-primate (cat), a strepsirhine primate (mouse lemur) and a haplorhine primate (gibbon). Animals received an intraocular injection of radioactive amino acids to reveal the distribution of retinal fibres and terminals. Drawings, scaled to the same size, are from coronal sections (lower right: approximate section levels shown on brain of gibbon) showing the relative positions of the three terminal nuclei (MTN, LTN, DTN, dark shading) of the AOS. Arrows, course of the superior fasciculus (SF-pf) (the main component of the AOT). As shown in the simplified schema in the upper left, this fasciculus separates from the main optic tract in the region of the brachium of the superior colliculus and then courses rostrally. The AOS in strepsirhine primates is similar to that of other mammals, but the MTNv is reduced in size. In haplorhine primates a retinal projection to the MTNv is lacking, and AOS fibres show a shortened trajectory through the cerebral peduncle to the MTNd. The extensive rostral development of the pons and pyramidal tracts in primates suggests that the reduction of the MTNv in Strepsirhini and its absence in Haplorhini is the consequence of a resulting morphological reorganization in the anterior region of the midbrain. AOS, accessory optic system; CP, cerebral peduncle; dLGN, dorsal part of the lateral geniculate nucleus; DTN, dorsal terminal nucleus of the AOS; HB, habenular nucleus; IF, inferior fasciculus; IP, interpeduncular nucleus: LTN, lateral terminal nucleus of the AOS; MG, medial geniculate nucleus; MM, mamillary nucleus; MTNd, dorsal division of the medial terminal nucleus of the AOS; MTNv, ventral division of the medial terminal nucleus of the AOS; n3, oculomotor nerve; N3, oculomotor nucleus; NOT, nucleus of the optic tract; NR, red nucleus; ON, olivary pretectal nucleus; PAG, periaqueductal gray; PRT, pretectum; SC, superior colliculus; SF-pf, posterior fibre branch of the superior fasciculus; SNc, substantia nigra, pars compacta; SNd, substantia nigra, pars diffusa.

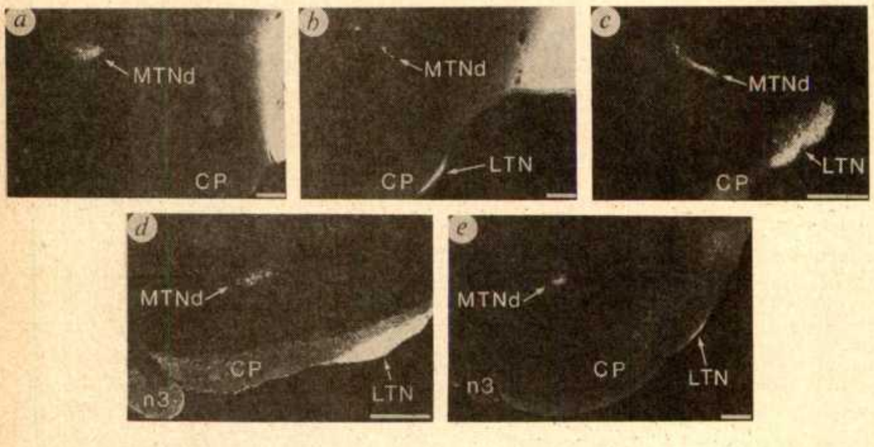


Fig. 2 Distribution of anterograde label transported from retinal ganglion cells to the MTNd (and the LTN) in: *a*, gibbon; *b*, macaque; *c*, marmoset; *d*, lemur; *e*, cat. The microphotographs of these coronal sections are viewed under dark-field illumination. Note the similar topographic location of the MTNd medial to the cerebral peduncle and the substantia nigra (towards upper left) in all species. Scale, 500 μm. CP, cerebral peduncle; LTN, lateral terminal nucleus of the accessory optic system; MTNd, dorsal division of the medial terminal nucleus; n3, oculomotor nerve.

1,000 μCi of ^{3}H-proline and ^{3}H-leucine) in the posterior vitreal chamber of the eye. After survival times of 24–60 hours the brains were processed according to the usual autoradiographic procedures for frozen or paraffin sections. In other primates, the eye was injected with 10–20 μl of a 40% solution of the anterograde tracer horseradish peroxidase (HRP; Sigma type VI) dissolved in sterile saline with 2% DMSO. Survival time was 36–60 hours. Frozen sections of the brains from these animals were treated with the tetramethylbenzidine method[10]. Brains from each series were sectioned in either frontal, sagittal, or horizontal planes and stained with cresyl violet to study cytoarchitecture. Retinal projections to the AOS were compared in species from eight other orders of mammals: Carnivora, Rodentia, Chiroptera, Dermoptera, Hyracoidea, Pholidota, Marsupiala and Edentata.

In all mammalian species studied, fibre tracts and the three terminal nuclei of the AOS were labelled by anterograde transport of the neuronal tracer from retinal ganglion cells. The majority of retinal fibres that form the accessory optic tract (AOT) cross in the optic chiasma, although a small component remains uncrossed. The AOT diverges from the main optic tract at the level of the brachium of the superior colliculus. The tract becomes visible on the superficial aspect of the brainstem as it retraces a rostral pathway to the anterior base of the midbrain[11], terminating in each of the AOS nuclei along its course (Fig. 1).

The MTN is the most conspicuous of the three AOS nuclei in non-primates and can be divided into a ventral division (MTNv) located at the ventromedial base of the cerebral peduncle, and a dorsal division (MTNd) extending into the region of the substantia nigra[3,12]. In primates, a retinal projection to the ventral division was observed only in Strepsirhines but the dorsal division receives retinal terminals in all species studied.

The projection to the MTNd is remarkably similar in primates and non-primates (Fig. 2). It is located in a similar position in all species (medial to the substantia nigra pars compacta and ventro-lateral to the medial lemniscus and the red nucleus). The distribution of terminal label (silver grains or HRP granules) in the MTNd is often reticular in appearance, as synaptic contacts of retinal axons are mainly situated on dendrites rather than on the cell bodies[13]. Fibres of the AOT course to the nucleus by penetrating directly through the cerebral peduncle, forming a characteristic obliquely-oriented fibre bundle in the region of the substantia nigra.

In some primates (*Macaca*, *Callithrix*) the MTNd is rather small, whereas in the gibbon (*Hylobates*) the nucleus extends rostrocaudally in the midbrain for >2.0 mm. The MTNd is partially embedded within the substantia nigra and in some instances the cells of these two structures actually intermingle. Several cytological features allowed us to distinguish the cells of each nucleus, however. The MTNd neurons are typically pale stained, and ovoid or fusiform in shape. In contrast, nigral cells are comparatively large sized, polygonal shaped, and intensely stained with cresyl violet; they also have a large clear cell nucleus and a dense compact nucleolus. Depending on the species, the two populations of cells are also distinguished by additional features. For example in the gibbon, as in other apes and man, the cytoplasm of nigral cells is densely packed with black granules of melanin[14]. These cells show an endogenous red fluorescence under dark-field illumination due to the presence of melanin whereas the MTNd cells, which lack pigment, do not fluoresce. Anterograde label was distributed only over the cells of the MTNd.

In the strepsirhine primates *Microcebus*[15], *Galago*[15] and *Nycticebus*[16], there is also a small but distinct ventral division of the MTN, which receives retinal afferents. As in other mammals, such as the cat, the MTNv is located at the base of the cerebral peduncle, immediately rostral to the emergence of the oculomotor nerve. The cells in this part of the MTN are densely stained, small and round in shape (average diameter 14 μm), and thus differ from those of the dorsal division, which are highly elongated (typically 8 μm × 40 μm). The cytoarchitecture and pattern of retinal projections to the two subdivisions of the MTN in *Microcebus* closely resemble those of the cat. In the cat the ventral and dorsal subdivisions clearly form a topographically continuous nucleus, whereas in *Microcebus* these two parts are interconnected by a few fine fibres. In haplorhine primates we could not identify either terminal label or a cell group typical of the MTNv.

The absence of a distinct ventral portion of MTN in haplorhine primates has been interpreted as the result of a simple regression of the nucleus during primate evolution[17,18]. However, we suggest that the neurons of the MTNv have migrated in position to a more dorsal location and form a single constituent nucleus with the MTNd. Such a dorsal shift of the MTNv in primates could be a consequence of the extensive development of the pyramidal tract[19] (related to precise motor control of the limbs), resulting in a rostral extension of the pons below the cerebral peduncle which influences the topographical relations between the nuclear groups in this region. Since in other mammals the neurones of the ventral and dorsal divisions of the MTN differ in cytoarchitecture[12], efferent connections[7,8] and histochemical properties[12], the hypothesis of a dorsal topological shift of the MTNv in primates could be experimentally verified if the same distinctions exist in this single constituent of the MTN.

Our results unequivocally demonstrate that in primates at least one subdivision of the MTN is present and receives a monosynaptic retinal projection. This conclusion is supported by the presence of retinal terminals to a well defined cellular aggregate located in a homologous region of the midbrain tegmentum and by the typical course of accessory optic fibre tracts to the nucleus. Preliminary results[20] also provide evidence that, as in the rat[7,8], the rabbit[7,8] and the cat[21], the MTN in *Macaca* also sends an efferent projection to the vestibular nuclei. Thus the intrinsic organization of the AOS and its relations with other sensory systems are constant features in all mammalian species.

We thank C. Blakemore for helpful criticism of the text and N. Boyer, P. Giroud, M. Rouvière and S. Bello for technical assistance. This work was supported by grants from INSERM, the DGRST and the Ministère de l'Education Nationale.

Received 24 January; accepted 6 October 1986.

1. Gudden, B. *Arch. Psychiat.* **2,** 364-366 (1870).
2. Fite, K. V., Brecha, N., Karten, H. J. & Hunt, S. P. *J. comp. Neurol.* **195,** 279-288 (1981).
3. Hayhow, W. R. *J. comp. Neurol.* **113,** 281-314 (1959).
4. Hayhow, W. R. *J. comp. Neurol.* **126,** 653-672 (1966).
5. Simpson. J. I., Soodack, R. E. & Hess, R. in *Reflex Control of Posture and Movement* (eds Granit, R. & Pompeiano, O.) 715-724 (Elsevier, Amsterdam, 1979).
6. Simpson, J. I. *A. Rev. Neurosci.* **7,** 13-41 (1984).
7. Giolli, R. A., Blanks, R. H. I. & Torigoe, Y. *J. comp. Neurol.* **227,** 228-251 (1984).
8. Giolli, R. A., Blanks, R. H. I., Torigoe, Y. & Williams, D. D. *J. comp. Neurol.* **232,** 99-116 (1985).
9. Pocock, R. I. *Proc. zool. Soc. Lond.* 19-55 (1918).
10. Mesulam, M. M. *J. Histochem. Cytochem.* **26,** 106-117 (1978).
11. Retzius, G. *Biol. Unters. Neue Folge* **8,** 65-74 (1898).
12. Gregory, K. M. & Giolli, R. A. *Expl Brain Res.* **60,** 501-508 (1985).
13. Lenn, N. J. *Brain Res.* **43,** 622-628 (1972).
14. Schwyn, R. C. & Fox, C. A. *J. Hirnforsch.* **15,** 95-126 (1974).
15. Cooper, H. M. *J. comp. Neurol.* **249,** 28-47 (1986).
16. Weber, J. T. & Giolli, R. A. *Brain Res.* **365,** 164-168 (1986).
17. Giolli, R. A. *J. comp. Neurol.* **121,** 89-108 (1963).
18. Lin, H. & Giolli, R. A. *Expl Brain Res.* **63,** 163-176 (1979).
19. Heffner, R. & Masterton, B. *Brain Behav. Evol.* **12,** 161-200 (1975).
20. Magnin, M. *et al.*, manuscript in preparation.
21. Magnin, M., Courjon, J. H. & Flandrin, J. M. *Expl Brain Res.* **51,** 298-303 (1983).

A cell surface molecule distributed in a dorsoventral gradient in the perinatal rat retina

Martha Constantine-Paton, Andrew S. Blum*, Rosalia Mendez-Otero† & Colin J. Barnstable*

Department of Biology, Yale University, Kline Biology Tower,
P.O. Box 6666, New Haven, Connecticut 06511, USA
* Laboratory of Neurobiology, The Rockefeller University,
New York, New York 10021, USA

Brain topography may have its earliest expression as spatial gradients of molecules controlling the deposition of neurones and neuronal processes[1,2]. In the vertebrate visual system there is evidence that the stereotyped alignment of central retinal projections relies on an initial spatially organized distribution of molecules in both the retina and its central target nuclei[3-6]. We used an immunological approach to look for molecules that are so organized and produced a monoclonal antibody (JONES) which shows a pronounced dorsal to ventral gradient of binding in the rat retina throughout the period when retinal ganglion cell axons are forming topographically organized projections within the central nervous system (CNS). Binding is present throughout the radial thickness of the retinal epithelium in regions where postmitotic neurones are generated but is not associated with any consistent histological characteristic of the tissue. The antibody was shown to bind on the cell surface of freshly dissociated retinal cells, and dorsal retinal quadrants were found *in vitro* to have nearly twice as much antigen as ventral retinal quadrants. Initial biochemical characterization of the target epitope reveals that it is a lipid present in chloroform/methanol extracts from perinatal retina and is sensitive to neuraminidase digestion.

---

† Permanent address: Universidade do Estado do Rio de Janeiro, Brazil.

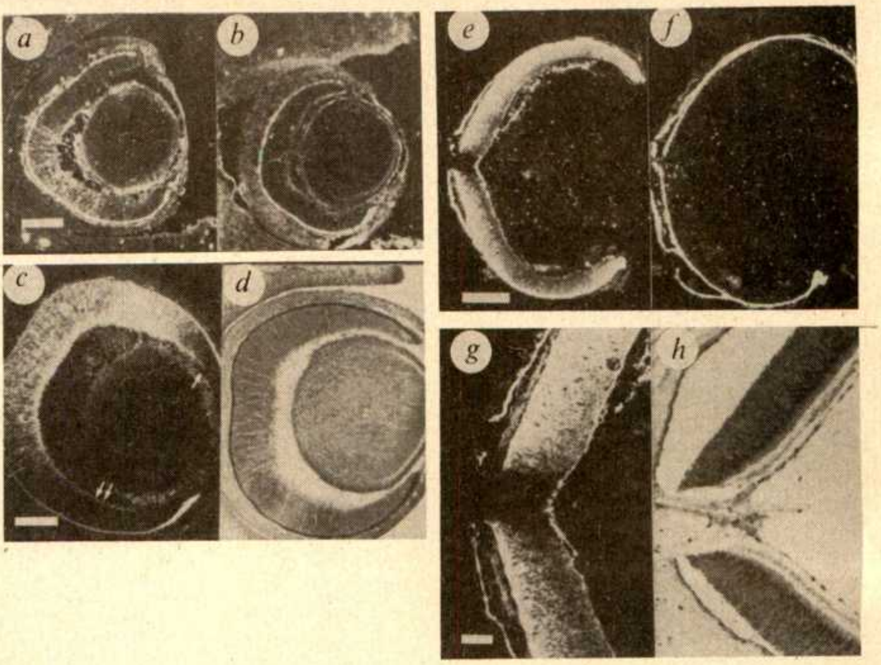


Fig. 1 Immunocytochemical distribution of JONES antibody binding in retinas sectioned parallel to their dorsoventral axes. *a*, Staining of central retina in the E13 eye; *b*, an adjacent 'blank' section reacted with second antibody alone; *c*, the dorsoventral gradient in the E17 eye. Arrow in dorsal retina, the initial expression of the antigen in the region adjacent to the zone at the ciliary margin that still contains only proliferating cells; double arrow, JONES binding in the optic fibre layer of both dorsal and ventral retina; *d*, an adjacent section counterstained with toluidine blue. Scale bars a–$d = 150\ \mu$m; *e*, the retinal epithelium at P0; JONES binding is now distributed throughout the epithelium but a dorsoventral gradient is still apparent; *f*, adjacent 'blank' section. Scale bar e, f $= 500\ \mu$m. In the P0–P3 retinas, the zone of neuroblast proliferation extends to the ciliary margin of the eye[7,8]. JONES binding also extends to the ciliary margin at these ages; *g*, higher magnification of the dorsoventral difference in JONES binding in the P0 eye. Horizontal sections of eyes at equivalent stages of development fail to show this pronounced anistropy in JONES binding; *h*, the equivalent region of another PO eye from a pup pulsed with ^{3}H-thymidine (1 μCi per g body wt., NEN, specific activity 6.7 Ci mmol^{-1}) two hours before it was killed. The most vitreal border of the proliferative zone in central retina is delineated by high grain densities over nuclei synthesizing DNA, post-mitotic amacrine and ganglion cells are present in inner retina. No differences between dorsal and ventral retina are apparent. Scale bar g, h $= 50\ \mu$m.

Methods. The JONES antibody was obtained by innoculation of a BALB/c mouse with retinal epithelia from fetal rats killed 14–21 d after sperm positive vaginal flushes, (E13–E20) rats. The mouse received three intraperitoneal injections of a crude homogenate of 6–10 lightly fixed retinas (2% paraformaldehyde for 30 min). For the first inoculation the retinal homogenates were suspended in complete Freund's adjuvant. Incomplete Freund's was used for the subsequent inoculations spaced two weeks apart. A final boost of ~5 retinas homogenized in phosphate buffered saline (PBS) was given intravenously four d before the fusion of spleen cells with P3-NSI/1-Ag4-1 plasmacytoma cells[33–35]. Screening of antibodies was performed with indirect immunofluorescence using tissue culture supernatants and sections of retina in which E13–E14 retinal epithelia had been nested inside P1 eyes. The JONES monoclonal antibody was sub-cloned three times by limiting dilution and shown by Ouchterlony analysis with subclass-specific sera to be immunoglobulin class M (IgM). Its binding has been examined using indirect immunofluorescence with goat anti-mouse IgG (H+L) (Cappel) or silver-intensified colloidal gold staining using a specific goat anti-mouse IgM (Janssen). Both procedures gave a staining pattern in the form of puncta that appeared to outline the perimeter of cell bodies. The retinal immunocytochemical staining above uses the silver intensified immunogold procedure in dark field illumination. Tissue was fixed overnight at 4 °C, cryoprotected in 30% sucrose in PBS and aligned in OCT mounting medium (Tissue-Tek, Miles) either as a whole fetal head or with the dorsal retinal pole indicated by a single remaining flap of skin. Sections were cut at 10 or 15 μm on a Hacker cryostat.

The JONES antigen was first detected in the rat at embryonic day 12 or 13 (E12–E13) where it was localized to central retina (Fig. 1*a*). By E17–E18 binding was observed along regions of the vitreal surface of the retina next to where an optic fibre layer had formed. In cellular regions of the epithelium the antigen was clearly distributed in gradient (Fig. 1*c*). JONES binding was invariably high in dorsal retina, decreasing gradually in more ventral retinal regions. In the E16–E17 rat a zone of postmitotic amacrine and ganglion cells is apparent in all but the peripheral rim of retina[7,8]. In dorsal retina the JONES antigen was expressed close to the boundary where these post-mitotic cells first appeared (Fig. 1*c*, *d*; arrows), but in ventral retina expression decreased to background levels well within the zone of post-mitotic ganglion and amacrine cells. The dorsal to ventral gradient of JONES binding was retained in the post-natal day 0 (P0) eye (Fig. 1*e*), and was still apparent, though less pronounced, at P3. Throughout the E17–P3 period ventral retina was indistinguishable from its dorsal counterpart on the basis of histological criteria or thymidine uptake (Fig. 1*d*, *g*, *h*).

Cells freshly dissociated from the neonatal eye expressed the JONES antigen on their surfaces (Fig. 2). Differences in JONES binding to different retinal quadrants were subsequently quantified using an *in vitro* assay (Fig. 3*a*). The small size of the early rodent eye precluded reliable dissections of the retina into subquadrant pieces and quantification of the absolute amount of binding was difficult because the degree to which the antibody penetrated the full thickness of the epithelium was unknown. Consequently the results of this assay do not give the maximal differences in antigen concentration between dorsal and ventral retina, but even so JONES binding was found to be ~1.5× higher in dorsal than in ventral retinal quadrants (Fig. 3*b*). This shows that the immunocytochemical differences in JONES binding do not result from fixation artefacts or variations in tissue mass.

Binding of JONES to aldehyde-fixed sections of late fetal and neonatal retina was removed by pretreatment of the fixed tissue with organic solvents and detergents, suggesting that the molecule recognized was a lipid. No antibody binding was observed on Western blots of late fetal or neonatal retinal extracts. Figure 4*a* illustrates immunoblots from thin layer chromatographic (TLC) plates of chloroform/methanol extracts of P4 retinas run in a chloroform: methanol solvent system. When such extracts were partitioned into aqueous and organic phases, all the immunoreactivity was found in the aqueous phase as a discrete band that migrated slightly below ganglioside G_{M2}. A single band of immunoreactivity was also found when the TLC plates were developed in a 1-propanol: aqueous calcium chloride solvent system (data not shown). No immunoreactivity suggestive of binding to lipoproteins or phospholipids was observed at the origin or in the upper regions of these lanes. The immunoreactive band was eliminated by treatment of the retinal extracts with neuraminidase (Fig. 4*b*). These results imply that the JONES epitope is carried on a relatively simple polar lipid, probably a minor species of ganglioside.

Position-dependent cell surface differences between dorsal and ventral retinal cells are suggested by preferential adhesion

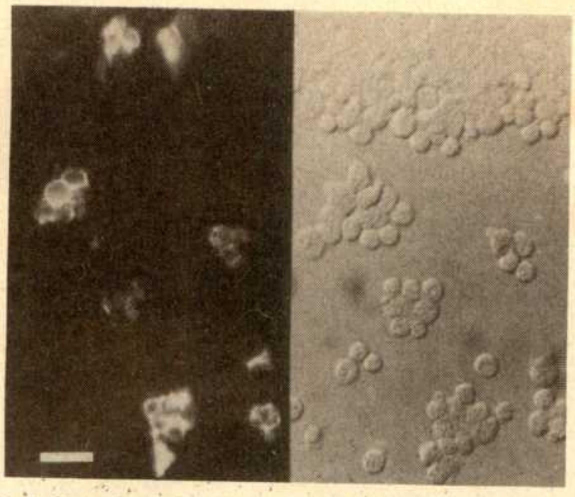
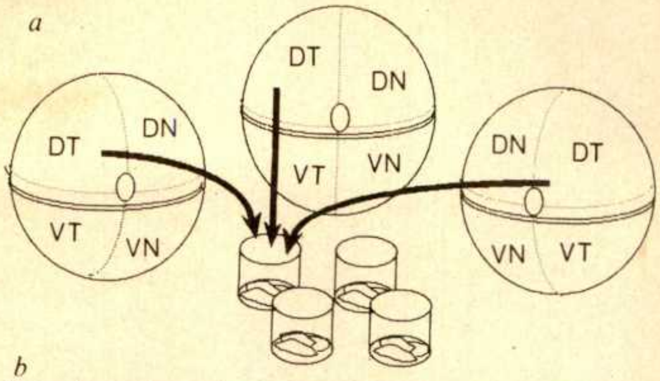


Fig. 2 Cells freshly dissociated from P2 retinal epithelia incubated with the JONES monoclonal antibody and then with rhodamine-labelled goat anti-mouse second antibody. Left, epifluorescent photomicrograph showing JONES labelling on the plasma membranes of cells; right, interference contrast photomicrograph of the same field. Scale bar = 20 μ.

Methods. Twenty P2 retinas were incubated for 20 min at 37 °C in 0.01% trypsin (Gibco 1:150), washed in PBS, titurated in PBS with DNase (5 μg ml^{-1}), centrifuged for 10 min, resuspended in 1 ml PBS and incubated for 1 h with the IgM fraction of JONES ascites fluid (1:100). The cells were then washed 3 times in PBS by centrifugation, resuspended in 1 ml PBS, incubated for 1 h at room temperature in goat anti-mouse IgG (1:50, H + L, Cappel), washed 3 times in PBS by centrifugation, pipetted onto a slide, covered and photographed immediately.



b

	DN c.p.m.	DT c.p.m.	VN c.p.m.	VT c.p.m.	D/V ratio	N/T ratio	DN/VT ratio	DT/VN ratio
1	3,393	5,315	2,856	2,799	1.54	0.77	1.21	1.86
2	5,877	3,818	3,451	3,217	1.45	1.33	1.83	1.11
3	3,918	2,710	2,674	2,130	1.38	1.36	1.84	1.01

Fig. 3 *a*, Diagrammatic representation of the assay for differential JONES binding to P3 retinal quadrants; *b*, the results of three replicates of this experiment using the quadrants from three retinas per assay point D, dorsal; V, ventral; N, nasal; T, temporal retinal halves. The difference between the summed counts for both dorsal and both ventral retinal quadrants was highly significant ($P < 0.0025$) by paired sample Student's *t*-test. Control experiments without JONES antibody or with other known non-reactive IgM antibodies revealed only low levels of binding and no differential binding to the retinal quadrants.

Methods. A large blood vessel running approximately horizontally along the back of the eye immediately ventral to the optic nerve head was used to distinguish dorsal from ventral retina. A cut perpendicular to this vessel running through the optic nerve distinguished nasal from temporal retinal halves. The pigment epithelium and inner limiting membrane were dissected from each retinal quadrant. The corresponding quadrants from three eyes were placed in one of 4 wells of a well culture plate containing 250 μl of minimal essential medium (MEM, Gibco), 10 μCi ^{3}H-leucine (specific activity 110 Ci mmol^{-1}) and 0.5 μl of ascites fluid from JONES hybridoma tumours. The retinal quadrants were incubated in this medium for 6 h at 37 °C, washed thoroughly in MEM and incubated for 1 h with ^{125}I-labelled F(ab)$_2$ fragments of an affinity-purified rabbit anti-mouse IgG directed against both heavy and light chains. The tissue was washed and processed for γ and β counting. ^{125}I counts were subsequently normalized to the amount of ^{3}H-leucine taken up and incorporated in each quadrant to control variations in the tissue volume in each well.

among cells from dorsal or ventral regions of chick retina[10]. Topographically appropriate binding of retinal cells to tectum was also observed[11], in experiments providing evidence for a double gradient where a protease-sensitive molecule controls the preferential adhesion of ventral retinal cells to dorsal tectum and a protease-insensitive molecule mediates the preferential adhesion of dorsal retinal cells to ventral tecta. The selectivity of binding of dorsal retinal cells was abolished by treatment with B-*N*-acetylhexosaminidase and could be mimicked by lecithin vesicles containing G_{M2} ganglioside[12]. A study using high performance liquid chromatography on lipid extracts of chick retina has recently demonstrated slightly higher endogeneous levels of G_{M2} in ventral retinal quadrants than in dorsal retinal quadrants[13]. The JONES antibody does not react with chick tissue. However, the initial biochemical characterization of the JONES antigen presented in this paper and studies using enzymatic treatments of extracts[14] suggest that it too may be a ganglioside. Monoclonal antibodies to ganglioside antigens (for example, 18B8[15,16], AbR24[17,18] and D1.1[19–21]) have been shown to vary in their cytoarchitectonic distribution with age in developing nervous system but no differences in binding based solely on topographic position have been described. To date, only one other molecule, TOP, has been found with a pronounced topographically varied distribution across the retina[22]. In binding assays the TOP molecule was nearly 12 times more concentrated in dorsal than ventral E14 chick retina[23]. However, it has not been possible to demonstrate a gradient of anti-TOP binding with immunocytochemistry. The JONES antigen is unlikely to be related to TOP since the TOP molecule appears to be glycoprotein[23].

JONES binding was observed in regions of the retinal epithelium shortly after the genesis of the first post-mitotic cells. In the early fetus this region is in the centre of the eye. In the chick[24] and frog eye[25,26] the first regions of the retina to have post-mitotic cells become situated in dorsal retinal quadrants because of the later addition of cells to ventral retina. Several observations suggest that similar displacements occur in the fetal rodent eye[27,28]. Thus the change from a symmetrical to a graded distribution of the JONES antigen between E13 and E17 could reflect a bulk tissue movement, continued expression by the earliest differentiating regions and a gradual spread of expression to ventral retina as the component cells reach the critical developmental stage. But this is unlikely to account for the continued expression of the gradient as, during the eight day period E17–P3 when the gradient of JONES binding is pronounced, even the youngest post-mitotic cells in the ventral regions of E17 retina attain the post-mitotic age and differentiate beyond the stage of dorsal retinal cells at E17 when the dorso-ventral gradient is first apparent. Therefore the differential expression of the JONES antigen is related to retinal position rather than to the post-mitotic age of individual retinal cells or to the differentiated state of a retinal region.

Molecules involved in establishing visual system topography are thought to be expressed in a graded pattern across the population of retinal ganglion cells so that they bias the selectivity of optic tract axons for different subregions of their central terminal fields[29]. However there is relatively little information on these molecules. The expression of JONES antigen in retina

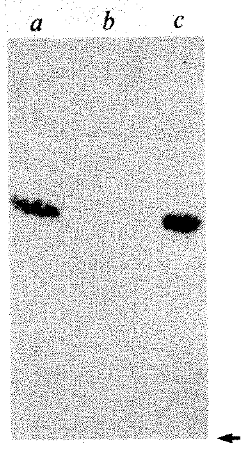


Fig. 4 Biochemical characterization of the JONES antigen using TLC/immunoblot analysis. *a*, The JONES antigen partitions into the aqueous phase of a Folch extract and is detected as a single band on immunoblots of lipids resolved by TLC. This band migrates to a position between that of G_{M2} and G_{M1} ganglioside standards run at the same time. (1) whole lipid extract. (2) organic phase; (3) aqueous phase; (4) G_{D1a}; (5) G_{M1}; (6) G_{M2}. 1-3, immunostaining of the nitrocellulose blot of a TLC plate; 4-6, resorcinol staining of purified standards run at the same time. *b*, The epitope recognized by JONES antibody contains sialic acid as it is removed by treatment with neuraminidase (+) but not buffer (−).

Methods. P4 rat retinas were homogenized in chloroform/methanol (2:1) to give a crude lipid extract[36]. This was partitioned into organic and aqueous phases by 0.1 M KCl[9]. Gangliosides and other highly polar lipids (including some phospholipids) partition into the aqueous phase, whereas most neutral glycolipids and some less polar gangliosides partition into the organic phase. The sialic acid content of the fractions was determined by the resorcinol method[37]. Samples containing 0.3-0.4 μg sialic acid were applied to a precoated HPTLC plate (Merck, silica gel 60, 0.2 mm thick) and developed with chloroform/methanol/0.25% aqueous $CaCl_2$ (60:40:10). Chromatograms were dried, wetted with isopropanol/water (2:1) and blotted onto nitrocellulose[38]. The nitrocellulose blot was blocked in 5% normal goat serum and incubated with JONES antibody for 4 h. After washing, the blot was incubated in peroxidase-conjugated goat anti-mouse IgG (H+L) for 2-4 h and washed. Bound antibody was visualized by incubation with 3.3′ diaminobenzidine. Ganglioside standards were run on parallel lanes as controls and were visualized by resorcinol staining[37]. For neuraminidase digestion the aqueous phase retinal lipids were dried under N_2 and resuspended in 0.1 M sodium acetate, *p*H 5.5, with or without neuraminidase (Sigma type X from *Clostridium perfringens*) at 8 mU per μg sialic acid. Incubations were carried out for 15 h at 37 °C. Digested lipids were extracted into chloroform:methanol, 2:1, dried under N_2 and resuspended in a small volume of chloroform:methanol, 2:1. Samples were run and analysed by TLC/immunoblot as described above.

coincides with the formation of central retinal connections. The first appearance of the JONES antigen in the retina on E12-13 is at the start of the period when the retinal ganglion cell axons first leave the eye[30]. These axons begin to terminate in the superior colliculus on E17 when the JONES gradient first becomes apparent in the retina. Throughout the period when the retinotectal map is laid down, the JONES antigen maintains its graded distribution in the cellular layers of retina and can also be localized to the retinal regions of optic fibre growth. Moreover, recent evidence indicates that the JONES antigen is also present in all the diencephalic and midbrain regions invaded by the retinal axons[31,32].

We thank Dr Burkhard Schlosshauer for help in the maintenance of the JONES hybridoma cell line, the production of the purified IgM antibody used for immunogold staining and immunofluorescent staining of freshly dissociated retinal cells. Supported by a John Simeon Guggenheim Fellowship to M.C-P. Biomedical Research Support Grant R07015 to Yale University, a Fulbright/Capes Fellowship award to R.M.-O., an MSTP Grant GM07739 award to A.S.B., NIH grants NS20483 and EY05206 and a Sloan Foundation Fellowship to C.J.B.

Received 12 May; accepted 25 September 1986.

1. Jacobson, M. *Developmental Neurobiology* (Plenum, New York, 1978).
2. Buskirk, D. R., Thiery, J-P., Rutishauser, U. & Edelman, G. M. *Nature* **285,** 488-489 (1980).
3. Gaze, R. M., Keating, M. J., Ostberg, A. & Chung, S.-H. *J. Embryol. exp. Morph.* **53,** 39-66 (1979).
4. Constantine-Paton, M. in *Developmental Order: Its origin and Regulation* (eds Subtelny, S. & Green, P. B.) 317-349 (Liss, New York, 1982).
5. Holt, C. E. *J. Neurosci.* **4,** 1130-1152 (1984).
6. Bonhoeffer, F. & Huf, J. *EMBO J.* **1,** 427-431 (1982).
7. Braekvelt, C. R. & Hollenberg, M. J. *Am. J. Anat.* **102,** 281-302 (1970).
8. Barnstable, C. J., Hofstein, R. & Akagawa, K. *Devl Brain Res.* **20,** 286-290 (1985).
9. Suzuki, K. *et al. J. Neurochem.* **12,** 629-638 (1965).
10. Gottlieb, D., Rock, K. & Glaser, L. *Proc. natn. Acad. Sci. U.S.A.* **73,** 410-414 (1976).
11. Barbera, A. J. *Devl. Biol.* **46,** 167-191 (1975).
12. Marchase, R. B. *J. Cell Biol.* **75,** 237-257 (1977).
13. Irwin, L. N., Bremer, E. G., Irwin, C. C. & McCluer, R. H. *Devl Neurosci.* **7,** 239-246 (1985).
14. Blum, A. S. *et al. Soc. Neurosci. Abstr.* **11,** 1064 (1985).
15. Grunwald, G. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **82,** 4008-4012 (1985).
16. Dubois, C. *et al. J. biol. Chem.* **261,** 3826-3830 (1986).
17. Goldman, J. E., Hirano, M., Yu, R. K. & Seyfried, T. N. *J. Neuroimmun.* **7,** 179-192 (1984).
18. Rosner, H., Al-Aqtum, M. & Henke-Fahle, S. *Devl Brain Res.* **18,** 85-95 (1985).
19. Levine, J. M., Beasly, L. & Stallcup, W. B. *J. Neurosci.* **4,** 820-831 (1984).
20. Cheresh, D. A., Reisfeld, R. A. & Varlei, A. P. *Science* **225,** 844-846 (1984).
21. Levine, J. M., Beasly, I. & Stallcup, W. B. *Devl. Brain Res.* **27,** 211-222 (1986).
22. Trisler, G. D., Schneider, M. D. & Nirenburg, M. *Proc. natn. Acad. Sci. U.S.A.* **78,** 2145-2149 (1981).
23. Trisler, G. D., Schneider, M. D., Moskal, J. R. & Nirenberg, M. in *Monoclonal Antibodies to Neural Antigens* (eds McKay, R., Raff, M. C. & Reichardt, L. F.) 231-265 (Cold Spring Harbor, New York, 1981).
24. Rager, G. *Adv. anat. Embryol. Cell Biol.* **63,** 1-90 (1980).
25. Holt, C. *Nature* **287,** 850-852 (1980).
26. Reh, T. E. & Constantine-Paton, M. *Devl Brain Res.* **10,** 187-200 (1983).
27. Silver, J. *Devl Biol.* **49,** 487-495 (1976).
28. Silver, J. & Robb, R. M. *Devl Biol.* **68,** 175-190 (1976).
29. Sperry, R. W. *Proc. natn. Acad. Sci. U.S.A.* **50,** 703-710 (1963).
30. Lund, R. D. & Bunt, A. H. *J. comp. Neurol.* **165,** 247-274 (1976).
31. Mendez-Otero, R. *et al. Soc. Neurosci. Abstr.* **12,** 317 (1986).
32. Schlosshauer, B. *et al. Soc. Neurosci. Abstr.* **12,** 317 (1986).
33. Kohler, G. & Milstein, C. *Nature* **256,** 495-497 (1975).
34. Barnstable, C. J. *Nature* **286,** 231-235 (1980).
35. Barnstable, C. J. & Drager, U. C. *Neuroscience* **11,** 847-855 (1984).
36. Folch, J., Lees, M. & Sloan-Stanley, G. H. *J. biol. Chem.* **226,** 497-509 (1957).
37. Svennerholm, L. *Biochim. biophys. Acta* **24,** 604-611 (1957).
38. Towbin, H., Schenen, C., Ball, R., Braun, D. G. & Rosenfeld, G. *J. immun. Meth.* **72,** 471-479 (1984).

Differentiation state-dependent surface mobilities of two forms of the neural cell adhesion molecule

G. Elisabeth Pollerberg, Melitta Schachner & Jean Davoust*

Department of Neurobiology, University of Heidelberg,
Im Neuenheimer Feld 364, 6900 Heidelberg, FRG
* European Molecular Biology Laboratory,
Postfach 10.2209, 6900 Heidelberg, FRG

The neural cell adhesion molecule (N-CAM) has been implicated in morphogenetic events during formation of the nervous system[1,2]. Three forms of N-CAM exist, all glycoprotein chains, of relative molecular masses 180,000 (180K), 140K and 120K (N-CAM$_{180}$, N-CAM$_{140}$ and N-CAM$_{120}$) which are differentially expressed on neural cell types and during development[3–5]. The three chains are thought to carry similar if not identical amino-acid sequences on their extracellular amino-terminal domains, but differ in the length of their carboxy-terminal cytoplasmic region[6–8]. They occur in highly sialylated embryonic and less sialylated adult forms[9,10]. N-CAM$_{180}$ is selectively expressed in more differentiated neural cells and may play a role in the stabilization of cell contacts[5]. To investigate this, we have studied in the surface membrane of a mouse neuroblastoma cell line N2A the lateral mobility of the two predominant forms of N-CAM, N-CAM$_{180}$ and N-CAM$_{140}$, as a function of differentiation. Here we report that as judged by fringe pattern photobleaching, the surface mobility of N-CAM$_{140}$ is higher than that of N-CAM$_{180}$, suggesting an association of N-CAM$_{180}$ with the cytoskeleton or other stabilizing factors. We also show that brain spectrin, a membrane-cytoskeleton linker

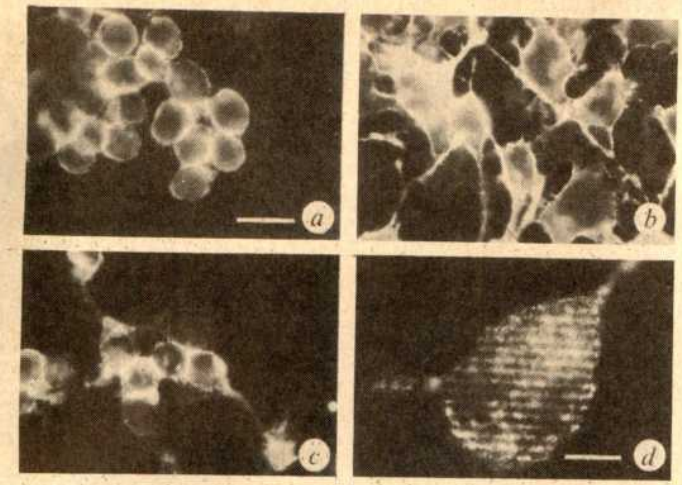


Fig. 1 Immunofluorescence localization of N-CAM in monolayer cultures of the mouse C1300 neuroblastoma clone N2A (ref. 35). Cells were stained with Fab fragments[36] of monoclonal antibody to mouse N-CAM (clone H28.123)[37,38] coupled to rhodamine (lissamine rhodamine B sulphonylchloride, Kodak)[39]. *a*, N2A cells maintained on coverslips coated with poly-L-lysine for 1 day in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS). *b*, N2A cells maintained for 1 day *in vitro* on coverslip coated with laminin (5 μg per ml DMEM for 12 h at 4 °C) in DMEM supplemented with 0.2% FCS. *c*, N2A cells cultured on coverslip coated with laminin for 7 day *in vitro* in DMEM supplemented with 0.2% FCS. *d*, Cells maintained as in *c* and illuminated with a fringe pattern. Scale bars, 30 μm (*a*, *b*, *c*) and 10 μm (*d*).

Methods. Cells were maintained under exponential growth conditions in plastic tissue culture flasks and subconfluent cells were trypsinized in the presence of EDTA as described[24], before being seeded at a density of 1×10^5 cells per ml on round 18 mm diameter glass coverslips under the culture conditions stated above. Live cells were incubated for 20 min at 0 °C with rhodamine-labelled Fab fragments of monoclonal N-CAM antibody (30 μg ml^{-1}) in Dulbecco's phosphate buffered saline (*p*H 7.3). They were rinsed 4 times with ice cold Hank's balanced salt solution, *p*H 7.3, supplemented with 0.1% bovine serum albumin. The coverslips were then subjected to fluorescence microscopy (*a*–*c*) after fixation with paraformaldehyde[40] or to photobleaching (*d*). For photobleaching, they were placed in a circular Plexiglass chamber containing 400 μl of the same medium. The chamber was quickly sealed on both faces with molten paraffin and transferred to the cold stage of the photobleaching microscope[13,14]. The cultures were maintained at 4 °C and a field diameter of 28 μm was illuminated by two coherent laser beams that create a fringe pattern with an interfringe spacing of 1.85 μm.

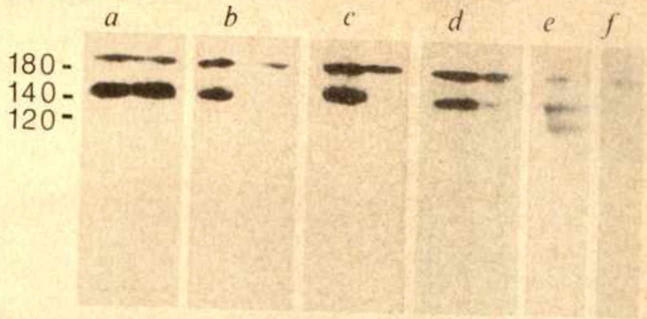


Fig. 2 Western blot analysis of: *a*–*d*, N-CAM expression by N2A cells maintained *in vitro* for 1 and 4 days (left and right lanes, respectively); *e*, immunoaffinity purified N-CAM; and *f*, radiolabelled brain spectrin bound to N-CAM. *a*, Cells seeded on poly-L-lysine-coated plastic flasks in DMEM supplemented with 10% FCS. *b*, Cells seeded on laminin-coated plastic flasks (see Fig. 1) in DMEM supplemented with 0.2% FCS. *c*, Cells plated on poly-L-lysine-coated plastic flasks in DMEM supplemented with 0.2% FCS and 2% dimethyl sulphoxide. *d*, Cells seeded on poly-L-lysine-coated plastic flasks in DMEM supplemented with 0.2% FCS. N2A cells were homogenized and further dissociated by ultrasonification (5 × 3 s). NP40 was then added to a final concentration of 0.5%. The homogenate was stirred for 30 min on ice, clarified by centrifugation for 3 min at 200*g* (Biofuge) and immediately boiled in sample buffer, separated by SDS-PAGE (7% slab gels), transferred to nitrocellulose filters, incubated with polyclonal antibodies to mouse N-CAM followed by incubation with ^{125}I-labelled protein A (2.5 μg ml^{-1}, 30 mCi mg^{-1}; Amersham) and autoradiographed[25]. *e*, Affinity-purified N-CAM from adult mouse brain[5] separated by SDS-PAGE was transferred to nitrocellulose filters, incubated with polyclonal rabbit antibodies to mouse N-CAM then horseradish peroxidase-coupled anti-rabbit Ig-antibody (Cappel) and developed with chloronaphtol as described[25]. *f*, Incubation of affinity-purified N-CAM (as in *e*) with ^{125}I-labelled brain spectrin[41] following the instructions of the supplier (Amersham) in a solid phase radioligand assay[42]. ^{125}I-labelled brain spectrin (specific activity 340,000 c.p.m. per μg) was used at a concentration of 1 μg ml^{-1} in 10 mM phosphate buffer, *p*H 7.4, containing 4% bovine serum albumin. Apparent relative molecular masses ($\times 10^{-3}$) are indicated on the left.

protein[11], binds only to N-CAM$_{180}$. The immobilization of N-CAM in differentiated N2A cells is achieved by a shift in expression from N-CAM$_{140}$ to N-CAM$_{180}$.

The mouse neuroblastoma cell line N2A can be maintained, depending on the culture conditions, in proliferating, morphologically undifferentiated and stationary (morphologically differentiated) states[12]. Under these conditions N2A cells only express the adult form of N-CAM. When cultured in low serum conditions or on laminin-coated coverslips, the cells stop proliferating and within one day of plating change from round cells with no processes to become more flattened, with numerous, sometimes branched neurites with growth cones, lamellipodia and filopodia (Fig. 1*b*). During the next 6–8 days, neurites become longer and less numerous (Fig. 1*c*). In low serum and with dimethyl sulphoxide, cell bodies assume a less flattened shape, but stop proliferating and extend neurites as on laminin. Because proliferation ceases under both conditions, differentiation does not alter the extent of cell contact except for the relatively small addition of the contacts made by outgrowing neurites. When laminin is used as substrate, N2A cells show a morphologically differentiated phenotype within a day (Fig. 1*b*), but express the two N-CAM components as in undifferentiated cells. It takes four more days for the gradual loss of N-CAM$_{140}$ (Fig. 2). When N2A cells are cultured on poly-L-lysine in dimethyl sulphoxide, the disappearance of N-CAM$_{140}$ correlates with morphological differentiation[5] (Fig. 2). We define undifferentiated cells as those that express both N-CAM$_{180}$ and N-CAM$_{140}$ and differentiated ones as those that only express N-CAM$_{180}$.

Lateral mobilities of cell surface constituents can be studied by determining the rate of diffusion of the fluorescence-labelled molecules from unphotobleached into photobleached areas. We created a series of stripes alternatively bleached and not bleached by using fringe pattern illumination (Fig. 1*d*). The decay in contrast of the periodic fluorescence concentration profile was measured as a function of time after pattern photobleaching[13,14]. To study the lateral mobility of N-CAM components, they were labelled using Fab fragments of a monoclonal antibody against N-CAM, directly conjugated with rhodamine. This antibody recognizes both forms of N-CAM[15,16], does not interfere with adhesion[16] and did not dissociate cell–cell or cell–substrate contacts in our experiments. As a control, we used rhodamine-labelled Fab fragments from polyclonal antibodies directed against mouse liver membranes. The antibodies react with an unrelated 130K antigen in undifferentiated and differentiated neuroblastoma cells (not shown). Single measurements were limited to a spot of 28 μm diameter, so that one cell and parts of neighbouring cells are photobleached. Data therefore represent the average N-CAM mobility in the surface membranes with and without cellular contact. This contact does not

Table 1 Lateral mobility of N-CAM and 130K antigen on surfaces of cultured neuroblastoma cells

Conditions of cell culture	*a*	*b*	*c*	*d*
Days in culture	1-3	1	7-9	7-9
DMEM supplemented with	10% FCS	0.2% FCS	0.2% FCS	0.2% FCS+2% DMSO
Coverslips coated with laminin	–	+	+	–
Morphological differentiation*	–	+	+	+
Expression of N-CAM components (kilodaltons)	180, 140	180, 140	180	180
Lateral mobility of N-CAM				
Mobile fraction† (%)	87 ± 2	87 ± 2	24 ± 3	17 ± 1
Diffusion coefficient† ($\times 10^{11}$ cm^2 s^{-1})	11.1 ± 1.1	14.7 ± 1.5	15.4 ± 2.4	10.3 ± 1.3
Number of determinations	26	26	10	23
Lateral mobility of 130K antigen				
Mobile fraction† (%)	91 ± 3	90 ± 2	83 ± 3	25 ± 1
Diffusion coefficient† ($\times 10^{11}$ cm^2 s^{-1})	8.9 ± 1.1	9.4 ± 1.5	13.1 ± 1.6	10.0 ± 1.4
Number of determinations	9	11	11	16

* Criteria for morphological differentiation are the following: undifferentiated, cells with round cell bodies and without processes; differentiated, cells with flattened cell bodies and several processes, sometimes branched, with growth cones and filopodia. DMSO, dimethyl sulphoxide.

† The fraction of mobile fluorochrome-coupled Fab fragments bound to live N2A cells was estimated from the modulated emission of fluorescence (MEF) curves (see Fig. 3) by fitting the data to a sum of two exponential components plus a constant baseline using a program developed by Vogel[34]. The decaying traces of lateral diffusion were found to contain essentially two terms: a rapid diffusion component in the range of 10^{-10} cm^2 s^{-1}, referred to here as the mobile fraction and a second component, either resulting from a very small diffusion coefficient, ($D < 4 \times 10^{-12}$ cm^2 s^{-1}) or due to subtle changes in the shape of the cell or even to a partial dissociation of Fab fragments. The diffusion coefficient reported here is given by $D = l^2/(4\pi^2 T)$, where l is the interfringe spacing (1.85 μm) and T the exponential decay time constant of the mobile fraction. The data are expressed as the mean ± s.d.

appear to be significant in determining the membrane mobility of N-CAM. Photobleaching measurements were performed at 4 °C to avoid internalization of antigen-antibody complexes, which was considerable with the 130K antigen and least significant with N-CAM$_{140}$.

After pattern photobleaching, the randomization of the immunolabelled N-CAMs occurred with two distinct diffusion kinetics (Fig. 3). In undifferentiated cells, the decay of the periodic fluorescence concentration profile is rapid (Table 1) and goes virtually to completion (Fig. 3*a*, curves *A*, *B*). Here, most of the N-CAM is of the N-CAM$_{140}$ form (see Fig. 2*a*, *b*) that has a high lateral mobility. In differentiated cells, there is a slow and incomplete randomization of N-CAM, characterized by a persistent periodic contrast (Fig. 3*a*, curve *C*). This immobile fraction of N-CAM is primarily N-CAM$_{180}$ which is exclusively present in differentiated cells (Fig. 2*c*). In contrast, the randomization of the 130K surface antigen is nearly complete during the time course of the experiment (Fig. 3*b*, curves *A*, *B*, *C*). The mobile fraction and respective mean diffusion coefficients were calculated (Table 1). N-CAM shows a highly mobile fraction which amounts to ~87% of all immunolabelled N-CAM molecules in the undifferentiated state and ~24% in the differentiated state of N2A cells. The diffusion coefficient is 1.1×10^{-10} cm^2 s^{-1} at 4 °C for undifferentiated cells, and approximately the same for differentiated cells. The 130K component has a mobile fraction amounting to approximately 91% of all immunolabelled 130K molecules in the undifferentiated state, 83% for cells induced to differentiate on laminin, and 25% on cells induced to differentiate in low serum and dimethyl sulphoxide for 7-9 days. Thus, the 130K component has a similar mobile fraction in undifferentiated and differentiated cells provided that differentiation is induced under the more physiological conditions of laminin as substrate. We cannot explain why cells allowed to differentiate in the presence of dimethyl sulphoxide contain a smaller fraction of mobile components. More importantly, however, the mean diffusion coefficient of the 130K component is nearly the same ($8.9\text{-}13.1 \times 10^{-11}$ cm^2 s^{-1}) in both the undifferentiated and differentiated states and does not depend on whether cells are induced to differentiate by dimethyl sulphoxide or by laminin. These experiments show a close correlation between the ratio of expression of N-CAM$_{140}$ to N-CAM$_{180}$ and the ratio of mobile to immobile N-CAM fractions in undifferentiated and differentiated states.

To test whether N-CAM$_{180}$ is linked to cytoskoletal components, immunoaffinity-purified N-CAM from adult mouse brain was separated by SDS polyacrylamide gel electrophoresis (SDS-PAGE), transferred to nitrocellulose and either stained with antibodies to N-CAM (Fig. 2*e*) or incubated with radiolabelled brain spectrin (Fig. 2*f*). Brain spectrin[11] binds exclusively to N-CAM$_{180}$.

We have shown that differentiation of neuroblastoma cells is accompanied by increased expression of N-CAM$_{180}$ which is less mobile in the surface membrane than N-CAM$_{140}$, the form expressed in undifferentiated cells. The rate of diffusion of N-CAM$_{140}$ is relatively high and similar to that of highly mobile cell surface glycoproteins (see, for instance, refs 17, 18). The immobilization of N-CAM is manifested by a decrease in the mobile N-CAM fraction (the N-CAM$_{140}$ component). This reduction in mobility is not due to a general differentiation-induced immobilization because another, unrelated membrane component of 130K has the same mobility in undifferentiated and differentiated cells. Morphological differentiation *per se* does not correlate with restriction in surface mobility or disappearance of N-CAM$_{140}$: when plated on the extracellular matrix component laminin, a physiological substrate for neurones[19-21], morphological differentiation of neuroblastoma cells is induced within one day. At this stage, both N-CAM$_{140}$ and N-CAM$_{180}$ are expressed and have a surface mobility characteristic of morphologically undifferentiated neuroblastoma cells. Morphological differentiation thus precedes N-CAM$_{180}$ immobilization. Failure to detect differences in surface mobilities between N-CAM and other cell surface molecules on neural cells[22] may have been due to an inability to distinguish between the individual N-CAM components.

The immobilization of N-CAM$_{180}$ in differentiated N2A cells could be related to an interaction with other cell surface components or with the cytoskeleton. We have previously shown that N-CAM$_{180}$ first appears in the central nervous system on post-mitotic neurones, whereas N-CAM$_{140}$ and N-CAM$_{120}$ appear at earlier developmental stages[5,23]. The appearance of N-CAM$_{180}$ is accompanied by the expression of another cell adhesion molecule, the L1 antigen[24-26], before the first migration of granule cell neurones in the cerebellar cortex[27]. N-CAM and L1 act synergistically in reaggregating cerebellar cells taken at the time of early postnatal granule cell migration[25]. L1 is associated with N-CAM$_{180}$, but not with N-CAM$_{140}$, as shown by

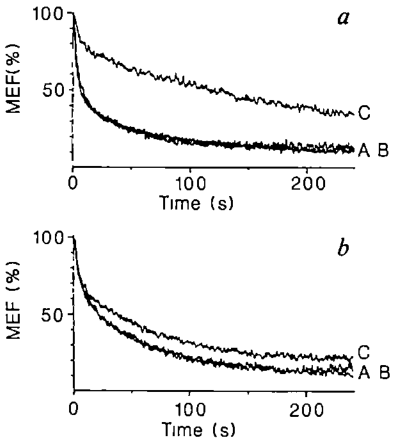


Fig 3 Modulated recovery of fluorescence after fringe pattern photobleaching of N2A cells *a*, Rhodamine-labelled Fab fragments of monoclonal N-CAM antibodies, *b*, rhodamine-labelled Fab fragments of polyclonal antibodies to a 130K antigen, *A*, N2A cells plated on poly-L-lysine in DMEM supplemented with 10% FCS and maintained in culture for 1-3 days *B*, N2A cells plated on laminin and maintained in DMEM supplemented with 0 2% FCS for 1 day in culture *C*, N2A cells plated on laminin and maintained in DMEM supplemented with 0 2% FCS for 7-9 days *in vitro*

Methods Polyclonal antibodies (raised in rabbits) against mouse liver membranes[27] recognize a single band of 130K relative molecular mass in Western blots of NP40 extracts of N2A cells (unpublished results) N2A cells were maintained on glass coverslips, labelled with fluorescent antibody and illuminated with a fringe pattern as described in Fig 1 Measurements were done on individual cells with or without contact sites During the bleaching pulse of 200 ms (1,500 W cm^{-2}, $\lambda = 514$ nm), between 10-20% of the total rhodamine-labelled Fab fragments present on the illuminated fringes with an interfringe spacing of 1 85 μm (see Fig 1*d*) are irreversibly bleached at the origin of the time axis This low level of bleaching was used to minimize photodamage to the cells After pattern photobleaching, the decay of the periodic fluorescence concentration profile was monitored continuously by an interrogation fringe pattern (less than 1 W cm^{-2}) that scans back and forth (3 kHz)[13,14] The resulting modulated emission of fluorescence (MEF) was detected using a photomultiplier coupled to a lock-in amplifier, and stored after digitalization with a time resolution of 240 ms Only the periodic distribution of bleached versus non-bleached fluorophores contributes to the signal The amplitude of MEF which is a measure of the periodic contrast created by the fringe pattern photobleaching was normalized with respect to its peak value (100%) The MEF decayed as a function of time because of the randomization of fluorochrome-labelled antibodies Each trace represents the average value of several measurements (n) $n = 26$ (*a*, *A* and *B*), $n = 10$ (*a*, *C*), $n = 9$ (*b*, *A*) and $n = 11$ (*b*, *B* and *C*) (see Table 1) Per coverslip, 5-10 individual photobleaching measurements could be performed within 1 h of mounting The standard deviations of the population curves were smaller than the peak-to-peak level of the noise of the averaged curve

antibody-induced co-redistribution on the surface of neuroblastoma cells[28] Both L1 and N-CAM$_{180}$ are concentrated at contact points between neighbouring neuroblastoma cells, whereas N-CAM$_{140}$ shows a more uniform distribution in both normal and transformed neural cells[5,29] The membrane-cytoskeleton linker protein brain spectrin (for reviews, see refs 30, 31) is also more concentrated at these contact points and copurifies with N-CAM$_{180}$, but not with N-CAM$_{140}$ nor N-CAM$_{120}$ under the stringent conditions of immunoaffinity purification[32] Combining these observations with the fact that N-CAM$_{180}$ shows a restricted lateral mobility in differentiated neuroblastoma cells suggests a direct cytoskeletal involvement in the restriction of lateral mobility, as in other systems[17,30,31,33] The membrane-cytoskeleton linker protein brain spectrin does indeed bind exclusively to N-CAM$_{180}$ The association of N-CAM$_{180}$ with the cytoskeleton could be mediated by the longer cytoplasmic domain of N-CAM$_{180}$ Further stabilization of cell contacts could be achieved by conformational modifications in intracellular protein domains triggered by conformational changes on extracellularly domains and vice versa It will be interesting to determine whether the restriction of N-CAM$_{180}$ within the surface membrane is induced by the preceding accumulation of the cytoskeleton at contact sites between neighbouring cells Alternatively, N-CAM$_{180}$ could be induced to accumulate by L1 antigen or other components localized within the surface membrane of the same cell, or by an extracellular component(s) In any case, restriction of lateral mobility and accumulation at particular sites within the surface membrane would then lead to the stabilization of selective cell-cell contacts

The authors are grateful to K Burridge for a gift of brain spectrin, C Goridis for clone H28 123, R Vogel for help with computer analysis and J Gruenberg and K Simons for helpful comments on the manuscript Supported by Deutsche Forschungsgemeinschaft (SFB 317), Fonds der chemischen Industrie and Studienstiftung des Deutschen Volkes (G E P)

Note added in proof It has been recently reported[43,44] that phosphatidylinositol is involved in the membrane attachment of NCAM$_{120}$ Therefore we would expect the surface mobility of NCAM$_{120}$ to be identical or higher than the one reported here for NCAM$_{140}$

Received 10 June, accepted 3 October 1986

1 Edelman, G M *A Rev Neurosci* **7**, 339-377 (1984)
2 Rutishauser, U *Nature* **310**, 549-554 (1984)
3 Chuong, C M & Edelman, G M *J Neurosci* **4**, 2354-2368 (1984)
4 Keilhauer, G , Faissner, A & Schachner, M *Nature* **316**, 728-730 (1985)
5 Pollerberg, E G , Sadoul, R , Goridis, C & Schachner, M *J Cell Biol* **101**, 1921-1929 (1985)
6 Gennarini, G , Hirn, M , Deagostini Bazin, H & Goridis, C *Eur J Biochem* **142**, 65-73 (1984)
7 Cunningham, B A Hoffman, S , Rutishauser, U , Hemperley J J & Edelman, G M *Proc natn Acad Sci USA* **80**, 3116-3120 (1983)
8 Hemperly, J J , Murray, B A , Edelman, G M & Cunningham, B A *Proc natn Acad Sci USA* **83**, 3030-3041 (1986)
9 Finne, J , Finne, H , Deagostini Bazin, H & Goridis, C *Eur J Biochem* **142**, 482-487 (1983)
10 Hoffman, S *et al J biol Chem* **257**, 7720-7729 (1982)
11 Burridge, K , Kelly T & Mangeat, P *J Cell Biol* **95**, 478-486 (1982)
12 Kimhi, Y , Palfrey C , Spector, I , Barak, Y & Littauer, U Z *Proc natn Acad Sci USA* **73**, 462-466 (1976)
13 Davoust, J , Stricker, R & Wijnaendts van Resandt, R *Biophys J* (in the press)
14 Davoust, J , Devaux, P F & Leger, L *EMBO J* **1**, 1233-1238 (1982)
15 Gennarini, G , Rougon, G , Deagostini Bazin, H Hirn, M & Goridis, C *Eur J Biochem* **142**, 57-64 (1984)
16 Goridis, C *et al Cold Spring Harb Symp quant Biol* **48**, 527-538 (1983)
17 McCloskey, M & Poo, M -m *Int Rev Cytol* **87**, 19-81 (1984)
18 de Laat, S W , van der Saag, P T , Elson, E L & Schlessinger, J *Proc natn Acad Sci USA* **77**, 1526-1528 (1980)
19 Rogers, S L , Letourneau, P C , Palm, S L , McCarthy, J & Furcht, L T J *Devl Biol* **98**, 212-220 (1983)
20 Baron van Evercooren, A *et al J Neurosci Res* **8**, 179-193 (1982)
21 Manthorpe, M *et al J Cell Biol* **97**, 1882-1890 (1983)
22 Gall, W E & Edelman, G M *Science* **213**, 903-905 (1981)
23 Nagata, I & Schachner, M *Neurosci Lett* **63**, 153-158 (1986)
24 Rathjen, F G & Schachner, M *EMBO J* **3**, 1-10 (1984)
25 Faissner, A , Kruse, J , Goridis, C , Bock, E & Schachner M *EMBO J* **3**, 733-737 (1984)
26 Bock E Richter-Landsberg, C Faissner, A & Schachner, M *EMBO J* **4** 2764-2768 (1985)
27 Lindner, J , Rathjen F & Schachner, M *Nature* **305**, 427-430 (1983)
28 Thor, G , Pollerberg G E & Schachner, M *Neurosci Lett* **66**, 121-126 (1986)
29 Pollerberg, G E thesis, Univ Heidelberg (1986)
30 Geiger, B *Biochim biophys Acta* **737**, 305-341 (1983)
31 Goodman, S & Zagon I S *Brain Res Bull* **13** 813-832 (1984)
32 Pollerberg, G E , Burridge, K , Krebs, K E , Goodman, S R & Schachner, M *Cell* (submitted)
33 Horwitz, A , Duggan, D , Buck, C , Beckerle, M C & Burridge, K *Nature* **320**, 531-532 (1986)
34 Vogel, R H *SPLMOD Users Manual EMBL Technical Report DAO8* (European Molecular Biology Laboratory, Heidelberg, 1986)
35 Augusti Tocco, G & Sato, G *Proc natn Acad Sci USA* **64**, 311-314 (1969)
36 Porter, R R *Biochem J* **73**, 119-131 (1959)
37 Hirn, M , Pierres, M , Deagostini-Bazin, H , Hirsch, M & Goridis C *Brain Res* **214**, 433-439 (1981)
38 Hirn, M *et al Neuroscience* **7**, 239-250 (1982)
39 Brandtzaeg, P *Scand J Immun* **2**, 273-290 (1973)
40 Schnitzer, J & Schachner, M *J Neuroimmun* **1**, 429-456 (1981)
41 Bolton, A E & Hunter, W H *Biochem J* **133**, 529-538 (1973)
42 Davis, J Q & Bennett, V *J biol Chem* **258**, 7757-7766 (1983)
43 He, H T , Barbet, J , Chaix, J C & Goridis, C *EMBO J* **5**, 2489-2494 (1986)
44 Sadoul, K , Meyer, A , Low, M & Schachner, M *Neurosci Lett* (in the press)

Endfeet of retinal glial cells have higher densities of ion channels that mediate K$^+$ buffering

Helen Brew, Peter T. A. Gray,* Peter Mobbs & David Attwell

Department of Physiology, University College London,
Gower Street, London WC1E 6BT, UK
* Department of Pharmacology, School of Pharmacy,
Brunswick Square, London WC1N 1AX, UK

A major function of glial cells in the central nervous system is to buffer the extracellular potassium concentration, $[K^+]_o$. A local rise in $[K^+]_o$ causes potassium ions to enter glial cells, which have membranes that are highly permeable to K$^+$; potassium then leaves the glial cells at other locations where $[K^+]_o$ has not risen. We report here the first study of the individual ion channels mediating potassium buffering by glial cells. The patch-clamp technique was employed to record single channel currents in Muller cells, the radial glia of the vertebrate retina. These cells have 94% of their potassium conductance in an endfoot apposed to the vitreous humour[1], causing K$^+$ released from active retinal neurones to be buffered preferentially to the vitreous[2,3] Recordings from patches of endfoot and cell body membrane show that a single type of inward-rectifying K$^+$ channel mediates potassium buffering at both cell locations. The non-uniform density of K$^+$ conductance is due to a non-uniform distribution of one type of K$^+$ channel, rather than to the cell expressing high conductance channels at the endfoot and low conductance channels elsewhere on the cell.

Cell-attached patch recordings[4] were made from the cell body or terminal endfoot membrane of Muller cells isolated from axolotl retinae (Fig 1*a*) As a rule, cell body recordings revealed no channels or only one channel, while recordings from the terminal endfoot membrane showed five or more (up to 50) channels Occasionally, however, cell body patches contained up to four channels and endfoot patches sometimes showed less than five channels

Over the voltage range where single channel transitions were seen, at potentials hyperpolarized from the channel reversal potential, the open channel current-voltage relations were linear with a conductance that was higher when the pipette $[K^+]_o$ was higher (Fig 1*b*) The current-voltage relations of membrane patches (Fig 2) suggest that the channels exhibit inward rectification, that is, the channel conductance is much smaller at potentials more positive than the reversal potential than at potentials below it This inward rectification may reflect the properties of the open channel I-V relation, or the presence of voltage-dependent gating, too fast to be detected as discrete channel jumps, decreasing the channel open probability at depolarized potentials[5,6] Extrapolated reversal potentials for the channels change by approximately 52 mV for a ten-fold change in $[K^+]_o$, indicating that the channels are largely potassium-selective (Fig 3*a*) The conductance is approximately proportional to the square root of $[K^+]_o$ (Fig 3*b*), as found for inward rectifier channels in other preparations[5,7,8] The potassium-selectivity and the $[K^+]_o$-dependence of the single channel conductance are indistinguishable for endfoot and cell body channels

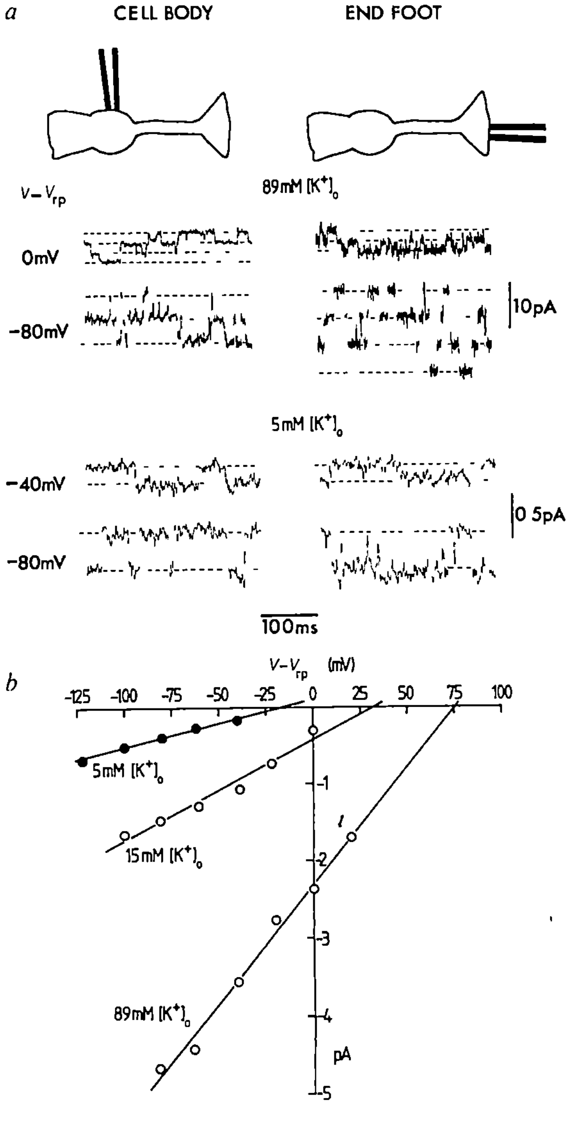


Fig. 1 *a*, Cell-attached recordings and *b*, single channel current-voltage plots for the cell body or terminal endfoot membrane of Muller cells isolated from enzymatically-dissociated[1] retinae of the axolotl (*Ambystoma mexicanum*) *a*, Records are shown from four different patches, two studied at each cell location, with 5 mM or 89 mM K$^+$ solution in the patch pipette (1 5 mM K$^+$ solution bathing the cell) Note the different current scales for the two $[K^+]_o$ values Inward current shown downwards For each patch, data are shown at the two potentials shown on the left relative to V_{rp}, the resting potential ($V-V_{rp}$ is minus the command potential applied to the pipette) Conventional micro-electrode recordings from 93 isolated Muller cells gave a mean resting potential of $V_{rp}=-88$ mV (±7 mV s d) Channel opening is downward (increased inward current) Dashed lines at constant current intervals show the single channel current levels A patch containing 3 channels has been chosen to display the 89 mM $[K^+]_o$ cell body data, typically there were one or no channels in cell body patches The K$^+$ channels characterized here were by far the most common channels observed They contributed the majority (typically 85% in 50 or 89 mM $[K^+]_o$) of the conductance in membrane patches studied at both the cell body and the endfoot the remaining conductance (not showing single channel transitions) was probably the conductance of the seal between the patch electrode and the cell, since its reversal potential was always around a pipette potential of zero rather than at the potassium reversal potential

Cells were maintained at 20-25 °C in perfusate containing (mM) NaCl 90 5, KCl 1 5, $MgCl_2$ 0 5, $CaCl_2$ 3, glucose 15, hydroxyethyl piperazine ethane sulphonic acid (HEPES) 20, adjusted to pH 7 3 with NaOH, bubbled with O_2 Solutions in the patch pipette were as above but with KCl replacing NaCl to give any desired $[K^+]_o$ value (5, 15, 30, 50 or 89 mM) *b*, Current-voltage plots obtained from records like those in *a*, for three patches studied with different concentrations of K$^+$ in the patch pipette Data for 5 mM $[K^+]_o$ are from an endfoot, 15 and 89 mM $[K^+]_o$ data are from cell body patches Each point plotted is the average of 10-20 single channel transitions Abscissa gives membrane potential relative to the resting potential in 1 5 mM K$^+$ solution (−88 mV)

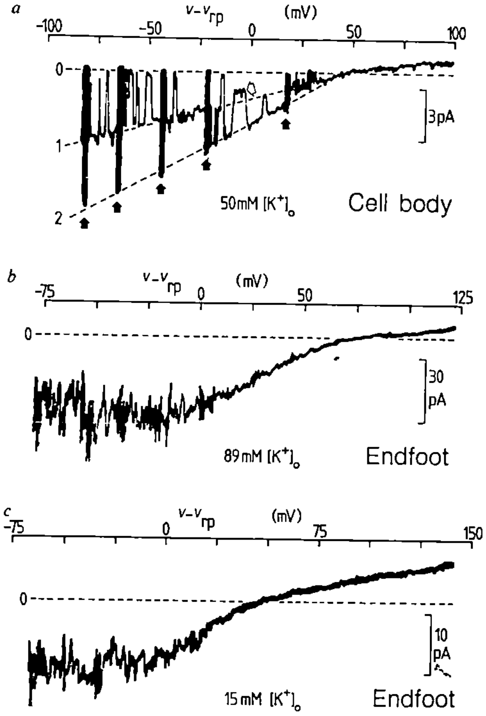


Fig. 2 Inward rectification of the Muller cell potassium channels. *a*, Current-voltage relation of a cell body membrane patch containing 2 channels, with 50 mM K$^+$ in the patch pipette. *b*,*c*, Current-voltage relations of endfoot patches showing inward rectification with 89 mM (*b*) or 15 mM (*c*) K$^+$ in the pipette. *a*, Current through the seal conductance (approximately 30% of the conductance of one open channel at negative potentials, with a reversal potential of 0 mV) was subtracted electronically from this record. The membrane potential (abscissa, given with respect to the resting potential of −88 mV) was ramped at roughly 70 mV s^{-1} from $V-V_{rp}=+100$ to −100 mV. Dashed lines connect values when no channels are open (0), when only one is open (1) and when both are open (2). At the voltages labelled with filled arrows, the voltage ramp was stopped to allow time for transitions to occur between all three states (0, 1 and 2). For $V-V_{rp}<-25$ mV, only 1 or 0 channels are open most of the time because of the low channel open probability, while for $V-V_{rp}>-25$ mV both channels are often open because the open probability is higher (see Fig. 4). For $V-V_{rp}>+50$ mV, above the channel reversal potential where the probability of channel opening is high (Fig. 4), the outward current carried by the channels is much less than would be predicted by extrapolating the channel I-V relation from more negative potentials. At one point, just negative to 0 mV, one channel enters a conductance substate (open arrow). Such substates were seen both at the endfoot and the cell body. At the cell body they were seen for 20% of channels and for these channels the probability of being in a substate was about 0.1 at −90 mV. *b*,*c*, Endfoot patches containing 24 and 30 channels respectively (calculated by dividing the inward current at $V-V_{rp}=0$ mV by the open probability at that potential (0.64, see Fig. 4) and by the single channel current). Flattening of the current for hyperpolarization below $V-V_{rp}=0$ mV reflects the voltage-dependence of the channel open probability (Fig. 4). For example, multiplying the curve describing the voltage-dependence of P_{open} in Fig. 4 by an ohmic I-V relation with a reversal potential of $V-V_{rp}=70$ mV predicts that the maximum inward current through the channels should be near $V-V_{rp}=-30$ mV and that the current at $V-V_{rp}=-80$ mV will be 84% of this maximum value, in reasonable agreement with the data in *b*. Above the reversal potential the outward current is much smaller than predicted by extrapolating from more negative potentials. No subtraction of the seal conductance was applied to records *b* and *c*.

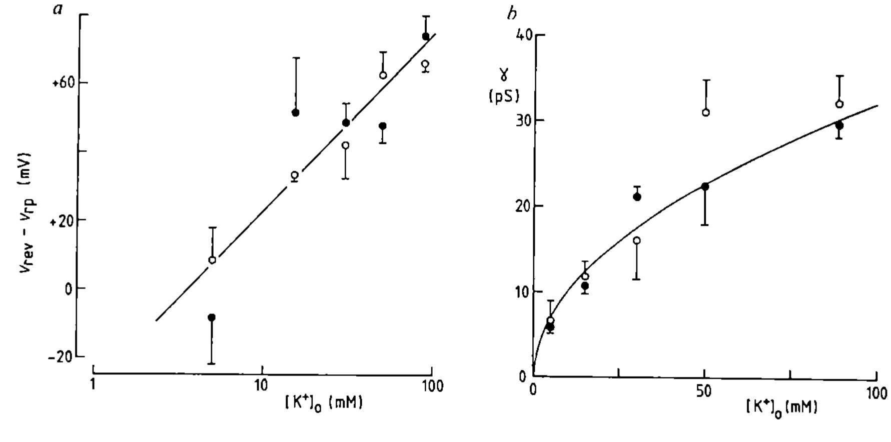


Fig. 3 *a*, Reversal potentials (measured relative to the resting potential of −88 mV) extrapolated from I-V plots like those in Fig. 1*b* for ○, cell body and ●, endfoot patches studied with 5 different pipette $[K^+]_o$ solutions. Each point is the average of 1-7 values, bars show standard errors. The line through the points is a least squares fit, with a slope of 52 mV per 10-fold change of $[K^+]_o$. *b*, Single channel conductance measured from I-V plots like those in Fig. 1*b*, for ○, cell body and ●, endfoot patches studied with 5 different $[K^+]_o$ values. Curve has the form $\gamma=(A[K^+]_o)^{\frac{1}{2}}$ with $A=10.24\ (\mathrm{pS})^2(\mathrm{mM})^{-1}$.

To investigate whether the probability of channels being open is higher at the endfoot than at the cell body, we constructed current amplitude histograms (Fig. 4*a*,*b*). Figure 4*c* and *d* shows the voltage-dependence of the steady-state open probability, calculated from such histograms, for endfoot and cell body channels. The open probability increases with depolarization, has similar values in endfoot and cell body channels and is independent of $[K^+]_o$ and $[Na^+]_o$.

We believe that the potassium channels characterized here are those responsible for mediating the spatial buffering of potassium by Muller cells for the following reasons. First, the channels have the necessary potassium specificity for carrying out potassium buffering. Second, the dependence of single channel conductance on potassium concentration in Fig. 3 is similar to that found for the conductance of whole Muller cells (measured at −90 mV by whole-cell clamping, for $[K^+]_o$ up to 20 mM, data not shown). Lastly, although Muller cell bodies contain I_A-type and calcium-activated potassium currents, in addition to inward-rectifying potassium currents, the former currents are only activated by depolarization beyond

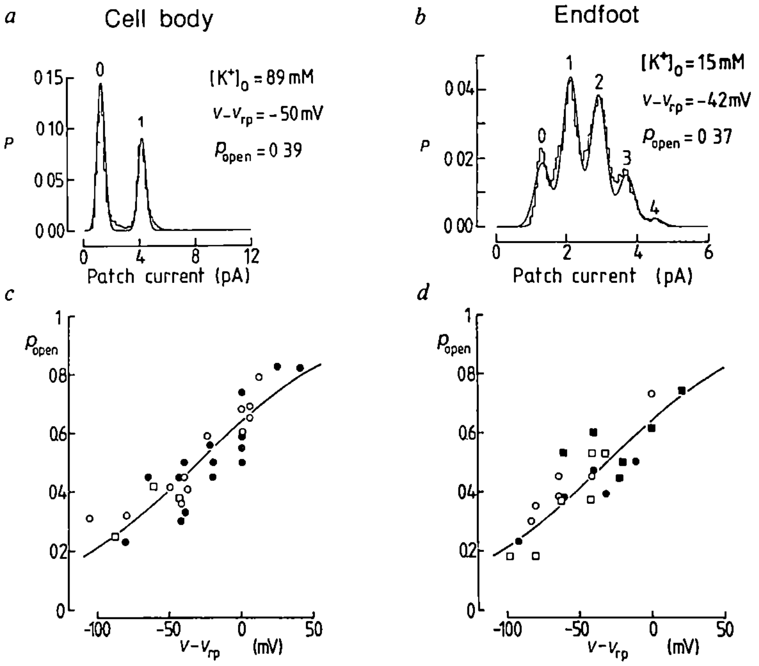


Fig. 4 *a,b*, Current amplitude histograms for determining the channel open probability, constructed from 10–40 second stretches of current at different potentials *c,d*, The open probability of cell body (*c*) and endfoot (*d*) channels, obtained at different voltages from data like those in *a* and *b*, with different values of $[K^+]_o$ ○, 89 mM K^+, ●, 50 mM K^+, ■, 30 mM K^+, □, 15 mM K^+ *a*, Cell body patch containing only one channel, held at a potential 50 mV hyperpolarized from the resting potential (at about −138 mV), 89 mM K^+ in the patch pipette P is the probability per 165 fA bin The two current levels (0 = no channels open, 1 = one channel open) are sufficiently separated on the amplitude histogram (by 2 95 pA) that the areas of their distribution peaks could be calculated directly The smooth curve through each peak shows a gaussian distribution with a standard deviation of 0 28 pA The open probability (area of 'open' peak/sum of areas of 2 peaks) was $P_{open} = 0\ 39$ *b*, An endfoot patch containing 4 channels, held 42 mV hyperpolarized from the resting potential (at about −130 mV), with 15 mM K^+ in the patch pipette P is the probability per 66 fA bin Curve through the points is a fit of the binomial prediction for the probability, $P(I)$, of a current amplitude, I, being seen $P(I) = \sum_{j=0}^{N} B(N,j)\, S(I - jI_1)$ where there are N channels present (4 in this case) each carrying an open channel current I_1, the current flowing when there are j channels open is jI_1, S is a gaussian function describing the noise present in the current (assumed, for simplicity to be independent of the number of channels open), $S(I) = \exp(-I^2/\sigma^2)$, and $B(N,j)$ is the probability of j of the N channels being open, obtained from the binomial prediction $B(N,j) = N!/[(N-j)!j!]P_{open}^{j}(1 - P_{open})^{N-j}$ where P_{open} is the open probability For these data $P_{open} = 0\ 37$, $I_1 = 0\ 81$ pA and $\sigma = 0\ 22$ pA *c,d*, Smooth curves through the points are best fits of the form $P_{open} = 1/\{1 + \exp(-(V - V_{rp} - V_{\frac{1}{2}})/k)\}$ The parameters $V_{\frac{1}{2}}$ and k were essentially identical at the cell body and endfoot $V_{\frac{1}{2}} = -30\ 5$ mV, $k = 53\ 2$ mV at the cell body, and $V_{\frac{1}{2}} = -31\ 4$ mV, $k = 54\ 9$ mV at the endfoot. This voltage-dependence of the open probability appears to be due to true voltage-dependent gating, rather than a voltage-dependent block by external Na^+ ions as occurs for the inward rectifier in tunicate eggs[8] and frog skeletal muscle[9], since (1) the dependence of P_{open} on voltage is independent of the extracellular sodium concentration (the variations of $[K^+]_o$ were obtained by substituting NaCl for KCl in the patch pipette) and (2) replacing extracellular NaCl by choline chloride did not change the open probability

−40 mV, which is probably more positive than physiological voltages[10]

This is the first study of the individual ion channels responsible for potassium buffering by glial cells It is also the first investigation attempting to account for the non-uniform distribution of K^+ conductance in glial cells in terms of the properties of individual ion channels Newman[10] suggested that potassium buffering might be mediated by different types of channels at the cell body and the endfoot We have shown that there is just one type of channel present, with more channels being expressed at the endfoot than elsewhere Recently, Newman[11] has shown that the non-uniform potassium conductance causing directed spatial buffering of potassium is not unique to Muller cells, but also occurs in optic nerve astrocytes The mechanisms we have characterized in Muller cells may therefore be common to all glial cells carrying out potassium buffering

Supported by the Wellcome Trust, the MRC and the Royal Society

Received 4 June, accepted 7 October 1986

1 Newman, E A *Nature* **309**, 155–157 (1984)
2 Newman, E A, Frambach, D A & Odette, L L *Science* **225**, 1174–1175 (1984)
3 Newman, E A & Odette, L L *J Neurophysiol* **51**, 164–182 (1984)
4 Hamill, O P, Marty, A, Neher, E, Sakmann, B & Sigworth, B *Pflugers Arch. ges Physiol.* **391**, 85–100 (1981)
5 Sakmann, B & Trube, G *J Physiol Lond.* **347**, 641–657 (1984)
6 Kurachi, Y *J Physiol Lond* **366**, 365–385 (1985)
7 Ohmori, H *J Physiol. Lond* **311**, 289–305 (1981)
8 Fukushima, Y *J Physiol Lond* **331**, 311–331 (1982)
9 Standen, N B & Stanfield, P R *J Physiol Lond* **294**, 497–520 (1979)
10 Newman, E A *Nature* **317**, 809–811 (1985)
11 Newman E A *Science* **233**, 453–454 (1986)

Cyclic GMP is involved in the excitation of invertebrate photoreceptors

Edwin C. Johnson, Phyllis R. Robinson & John E. Lisman

Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA

The hyperpolarizing receptor potential in vertebrate rod photoreceptors appears to be mediated by the second messenger, cyclic GMP. Injection of cGMP into rods[1,2] or application of cGMP to inside-out membrane patches[3] activates a conductance resembling that produced by light. Light produces a rapid reduction of cGMP in living rods[4], leading to closure of sodium channels and membrane hyperpolarization. In most invertebrate photoreceptors the response to light is depolarizing. We have investigated whether cGMP is involved in controlling the increase in sodium conductance that underlies this depolarization[5]. We show here that injection of cGMP into *Limulus* photoreceptors produces a depolarization that mimics the receptor potential. We also show that the cGMP concentration of the squid retina increases rapidly during exposure to light. These results support the hypothesis that cGMP mediates the light-induced depolarization in invertebrate photoreceptors and suggests that vertebrate and invertebrate phototransduction may be more similar than previously thought.

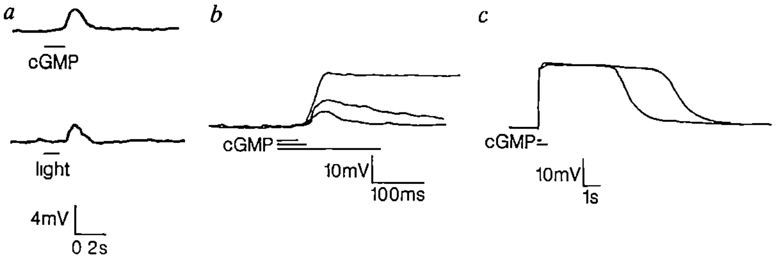


Fig. 1 Responses to injection of cGMP Individual photoreceptor cells were mechanically isolated from a desheathed *Limulus* ventral nerve with a large suction pipette in order to distinguish the light-sensitive lobe A microelectrode containing cGMP was placed in the R-lobe under the control of a motor-driven micromanipulator Injections were monitored optically with an infrared television[17] The microelectrode had 10 mM or 1 mM cGMP in stock solution of 300 mM K aspartate, 10 mM HEPES, *p*H 7 0 *a*, The response to a single photon (lower trace) compared with the response to a very small injection of cGMP (upper trace) *b*,*c*, Superimposed responses to pressure injection of increasing amounts of cGMP Duration of injections is marked by bars below the voltage traces The applied pressure was <1 p s i

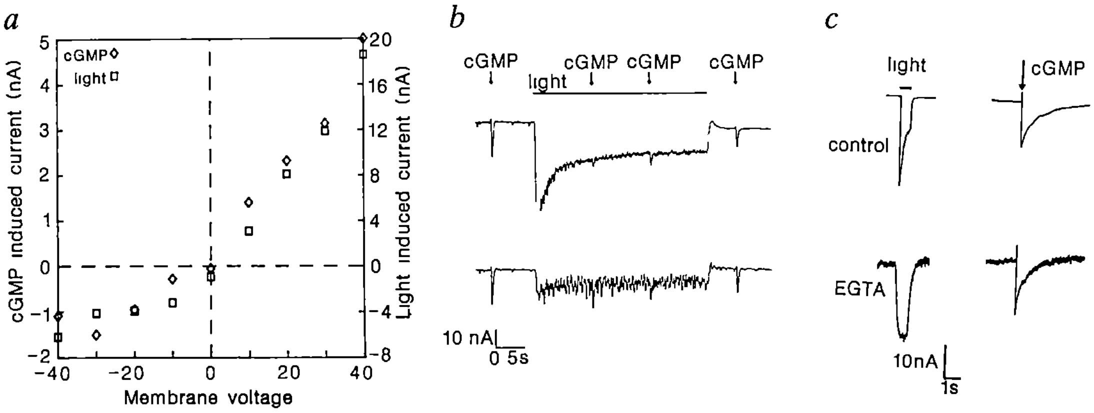


Fig 2 Effect of voltage, background light, and intracellular Ca^{2+} buffer on cGMP-induced currents measured using two-electrode voltage-clamp *a*, Current-voltage characteristics of light-induced and cGMP-induced currents Light intensity 0 08 μW cm^{-2}, currents are given in nA *b*, Effect of adapting lights on the response to cGMP A brief pulse of cGMP was injected at each arrow Upper trace light intensity 0 05 μW cm^{-2}, lower trace 0 005 μW cm^{-2} Note that the light-adapted cGMP response at the end of the upper trace recovered in the 30-s dark period between the upper and lower traces *c*, Response to cGMP was not blocked by injection of EGTA from a second electrode Note that this is a different cell and slower time base than in *b* The responses on the left demonstrate that the light response after EGTA injection (lower left) did not display the decline from the initial transient phase to the plateau phase normally observed in the light response (upper left) This is evidence that EGTA buffered the change in intracellular calcium level that regulates light adaptation[18] On the right are responses to cGMP injection in the dark (top right, before EGTA injection, bottom right, after EGTA injection) The reason for the increase in baseline noise after EGTA injection is unclear

We pressure-injected cGMP into the cytoplasm of *Limulus* ventral photoreceptors If the injection electrode was in a cGMP-sensitive region of the cell (see below), injections evoked a transient depolarization For small injections, the response kinetics were similar to the kinetics of the light response (Fig 1*a*) The size of the depolarization was smoothly graded with the strength of the injection (Fig 1*b*), there was no indication that cGMP can evoke discrete waves[6], as is characteristic of other agents that depolarize ventral photoreceptors[7-9] With larger injections the response saturated, the only effect of further increases was to lengthen the response duration (Fig 1*c*) Once the response began to fall, it fell with kinetics that were independent of the duration of the response The saturation of the cGMP response was only 20–30 mV, much lower than the saturation of the light response (50–60 mV) This is as expected if the injected cGMP saturates a local subpopulation of channels The cGMP would then be rapidly hydrolysed, but the voltage would not fall until the local cGMP concentration fell below saturation The larger the injection, the longer cGMP levels would remain at saturation Voltage-clamp measurements allowed further comparison of the responses to light and cGMP cGMP induced a current with a reversal potential and voltage-dependence nearly identical to that of the light-induced current (Fig 2*a*) The response to cGMP was reduced by an adapting light the same way as the response to a flash would be (Fig 2*b*) Thus, in several respects the conductance activated by cGMP is similar to that activated by light

Recently inositol trisphosphate ($InsP_3$) has been suggested as a messenger of excitation in invertebrate photoreceptors[7,8] One difficulty with this hypothesis is that injection of calcium buffers blocks the excitation produced by $InsP_3$ (ref 10) but enhances the maintained excitation produced by light[18] Injection of calcium buffer (EGTA) did not block the response of a dark-adapted cell to cGMP (Fig 2*c*) Thus both the response to light and cGMP can occur without a change in cytoplasmic calcium

Not all regions of the *Limulus* ventral photoreceptor were sensitive to cGMP The ventral photoreceptor is divided into lobes[11,12], the light-sensitive R-lobe (40 μm in diameter) contains rhodopsin in microvillar specializations of the plasma membrane The A-lobe contains no microvilli Microvilli are restricted to small subregions of the R-lobe The sensitivity of the cell to cGMP was highly dependent on the position of the pipette Injection of cGMP into the A-lobe was without effect Injection of cGMP into the R-lobe often produced only small depolarizations (0–5 mV) By moving the injection electrode through the R-lobe, it was sometimes (<10%) possible to find a region in which injections produced large responses The position dependence of the cGMP response was confirmed by making reversible movements in and out of the cGMP-sensitive region (Fig 3*a*) Because of this position dependence, it was difficult to study the cGMP response This probably explains the failure of previous efforts[9,13] to detect a cGMP response

Because of the low probability of obtaining responses to cGMP it was possible that these responses were artefacts It was therefore important to see if other substances produced responses when injected into regions known to be responsive to cGMP We used double-barrel θ micropipettes with one barrel containing cGMP, the other a test solution Once a cGMP-

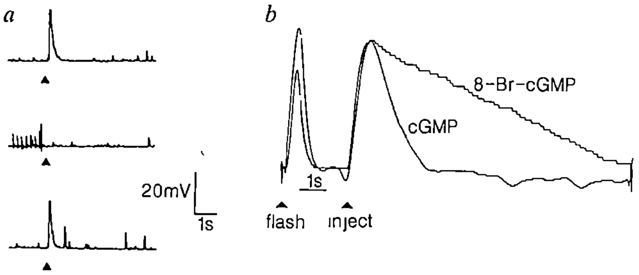


Fig. 3 Dependence of the response to cGMP on electrode placement and cGMP hydrolysis rate. *a*, Dependence of response to cGMP on electrode position. Upper trace, response to brief injection of cGMP (▲); middle trace, after ~10 μm movement of electrode (stepping motor caused series of artefacts) there was no response to cGMP injection; bottom trace, movement of micropipette back to original position restored the response. *b*, Response to the slowly hydrolysable analogue 8-Br-cGMP. A double-barrel pipette with 10 mM cGMP in one barrel and 10 mM 8-Br-cGMP in the other barrel was used. Injections given at arrows. Average responses from 7 injections of each are shown superimposed. Peak responses to 8-Br-cGMP and cGMP differed by <20%, but are shown here normalized. Depolarizations at left are responses to test flashes.

sensitive region was found, larger amounts of the test solution were injected through the other barrel. We found that cGMP-sensitive regions showed no significant response to the injection stock, cAMP or 5′-GMP, but were sensitive to 8-Br-cGMP. The response to 8-Br-cGMP fell more slowly than the response to cGMP (Fig. 3*b*) as would be expected of this slowly hydrolysable analogue[14,15].

The very rapid hydrolysis of cGMP must limit its spread and may explain why the response to cGMP is more localized than the response to Ca^{2+} (ref. 21) or $InsP_3$ (refs 7, 8). We suspect that localization may also be due to restricted access to the microvillar membrane, possibly caused by the subhadromeric reticulum, a fenestrated internal membrane system that abuts the microvillar membrane[10]. Indeed, a quasi-independent compartment between the reticular and microvillar membranes might enhance transduction, since a change in the number of cGMP molecules leads to a concentration change inversely proportional to the volume over which cGMP is spread; a restricted diffusion therefore enhances the signal.

If cGMP is a messenger for excitation in invertebrate photoreceptors, cGMP levels must rise rapidly in light. We examined cGMP levels in squid (*Loligo paeili*) because its retina is large and primarily contains photoreceptors. Punches ($\frac{1}{4}$ inch diameter) of living retina and associated scleral material were obtained from dark-adapted animals under deep-red illumination and were frozen with a liquid nitrogen cooled hammer. The punch from one eye was frozen in the dark, that from the other eye was frozen 100–200 ms after the onset of illumination. cGMP was extracted by homogenizing the frozen retinas in cold HCl-ethanol (1:99), and measured by radioimmunoassay[19]. Dark levels of cGMP were 2.0 ± 0.3 (s.e.m., $n=23$) pmol per mg protein. Light levels were 1.7 ± 0.27 (s.e.m., $n=23$) times greater than dark levels, a significant increase ($t=1.87$, d.f. $=22$, $P<0.05$, one tail). As this is the sum of cGMP changes over the whole punch of retina, the fractional change of cGMP in the transducing region is likely to be much larger than 1.7.

The finding of a light-induced increase in cGMP in the intact squid retina is consistent with the *in vitro* work of Saibil[16] showing a light-induced rise of cGMP in homogenates of squid outer segments. Our general conclusion regarding the importance of cGMP is consistent with the early suggestion[20] that cyclic nucleotides are involved in invertebrate phototransduction. Previous work on vertebrate rods[1–3] and the work reported here on invertebrate photoreceptors suggest that cGMP increases the sodium conductance in both types of cells. The difference in sign of the receptor potentials appears to be due to a difference in sign of the light-induced change in cGMP. The decrease in cGMP concentration in vertebrates is due to activation of phosphodiesterase. Whether the rise in cGMP concentration in invertebrates is due to activation of guanylate cyclase and/or inactivation of phosphodiesterase remains to be determined.

We thank Larry Pinto for advice on the use of double-barrel electrodes, Jim Nathanson for advice on preparation of frozen tissue for cGMP analysis, and Joel Brown, Gordon Fain and Irwin Levitan for critical readings of this manuscript. This work was supported by NIH grants EY01496 and EY05807.

Received 13 June; accepted 10 September 1986.

1. Miller, W. H. & Nichol, G. D. *Nature* **280**, 64–66 (1979).
2. MacLeish, P. R., Schwartz, E. A. & Tachibana, M. *J. Physiol., Lond.* **348**, 645–664 (1984).
3. Fesenko, E., Kolesnikov, S. S. & Lyubarsky, A. *Nature* **313**, 310–313 (1985).
4. Woodruff, M. L., Bownds, D., Green, S. H., Morrisey, J. L. & Shedlovski, A. *J. gen. Physiol.* **69**, 667–679 (1977).
5. Millechia, R. & Mauro, A. *J. gen. Physiol.* **54**, 331–351 (1969).
6. Fuortes, M. G. F. & Yeandle, S. *J. gen. Physiol.* **47**, 443–463 (1964).
7. Brown, J. E. *et al. Nature* **311**, 160–163 (1984).
8. Fein, A., Payne, R., Corson, D. W., Berridge, M. J. & Irvine, R. F. *Nature* **311**, 157–160 (1984).
9. Fein, A. & Corson, D. W. *Science* **204**, 77–79 (1979).
10. Payne, R., Corson, D. W., Fein, A. & Berridge, M. J. *J. gen. Physiol.* **88**, 127–142 (1986).
11. Calman, B. & Chamberlain, S. J. *J. gen. Physiol.* **80**, 839–862 (1982).
12. Stern, J., Chinn, K., Bacigalupo, J. & Lisman, J. *J. gen. Physiol.* **80**, 825–837 (1982).
13. Stern, J. & Lisman, J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7580–7584 (1982).
14. Robinson, P. R., Radeke, M. J., Cote, R. H. & Bownds, M. D. *J. biol. Chem.* **261**, 313–318 (1986).
15. Zimmerman, A. L., Yamanaka, G., Eckstein, F., Baylor, D. A. & Stryer, L. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8813–8817 (1985).
16. Saibil, H. *FEBS Lett.* **168**, 213–216 (1984).
17. Corson, D. W. & Fain, A. *Biophys. J.* **44**, 299–304 (1983).
18. Lisman, J. E. & Brown, J. E. *J. gen. Physiol.* **66**, 489–506 (1975).
19. Steiner, A. L. *Meth. Enzym.* **38**, 96–112 (1974).
20. Bitensky, M. W., Gorman, R. E. & Miller, W. H. *Proc. natn. Acad. Sci. U.S.A.* **68**, 561–526 (1971).
21. Payne, R., Corson, D. W. & Fein, A. *J. gen. Physiol.* **88**, 107–126 (1986).

Characterization of single pacemaker channels in cardiac sino-atrial node cells

Dario DiFrancesco

Università di Milano, Dipartimento di Fisiologia e Biochimica Generali, Elettrofisiologia, via Celoria 26, 20133 Milano, Italy

Normal pacemaking in the mammalian heart is driven by spontaneously active cells located in the sino-atrial (SA) node. The rate of firing of these cells and the modulation of this rate by catecholamines are controlled by i_f, an inward Na- and K-current that turns on at voltages more negative than −40 mV[1–9]. The 'pacemaker' current i_f is also present in other types of cell where its ability to produce and modulate a depolarizing process may be useful[10]. For example, in vertebrate photoreceptors i_f drives the depolarization that terminates the light-induced hyperpolarization[11]. Currents similar to i_f are also found in hippocampal neurones and DRG neurones[12,13]. The present report shows for the first time that the opening of single i_f-channels of low conductance (1 pS) can be resolved using a modification of the patch-clamp technique[14] on isolated SA-node cells. Modulation of i_f by adrenaline is shown to be mediated by an increase in the probability of channel opening, whereas the single-channel amplitude remains unchanged.

Although the 'pacemaker' current was one of the first currents to be discovered in cardiac tissues and has a mixed ionic nature which suggests a relatively large single-channel conductance such as that found for other non-specific cation currents[15–17], attempts at recording the activity of single i_f channels have previously been unsuccessful. This has led to the suggestion that either the elementary current steps are too small, or the current is carried by a transport mechanism that is not a channel[18].

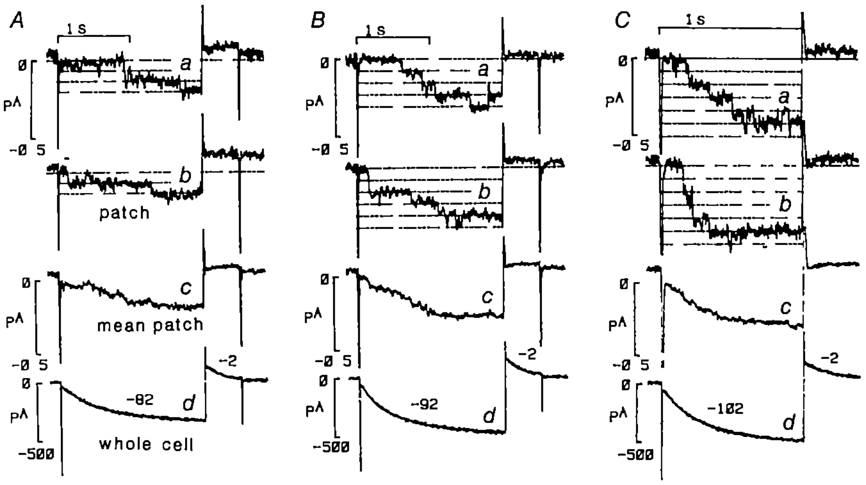


Fig. 1 Simultaneous measurements of whole-cell i_f and single i_f channels in a sino-atrial node cell. The bottom traces (labelled *d*) represent i_f recorded with a pipette in the whole-cell configuration during voltage-clamp steps from a holding potential of −32 mV. Voltages are labelled near corresponding traces. Activity from a membrane patch of the same cell was simultaneously recorded using an independent pipette in the cell-attached configuration. The top traces (*a* and *b*) are examples of patch recordings during i_f activation showing multiple step-like transitions. These transitions represent consecutive openings of single channels during hyperpolarization. Grids with fixed intervals were used to fit patch records by eye and to determine the elementary current step at each potential. These were −0.067, −0.077 and −0.085 pA at −82, −92 and −102 mV, respectively. The traces labelled *c* are averages of patch records ($n = 7$, 7 and 9 in *A*, *B* and *C*, respectively). Their time courses follow those of the corresponding whole-cell currents, indicating that patch activity reflects opening of single i_f channels. The maximum number of elementary openings was 3 at −82 mV, 5 at −92 mV and 6 at −102 mV. Notice that the time courses of mean-patch and whole-cell traces differ at −2 mV. This can be attributed to the different Na and K concentrations used for the patch-pipette and extracellular solutions (below) which determine different reversal potentials for the two traces. A similar observation applies to the data in Fig. 3.

Methods Cells were dissociated from the sino-atrial region of rabbit hearts as described[19]. Their mean resistance and capacity were 1213 ± 417 MΩ (number of recordings (n) = 17) and 29.3 ± 9.5 pF ($n = 20$), respectively. The cells were superfused with warm (35–36 °C) Tyrode solution containing (in mM): NaCl 140, KCl 5.4, $MgCl_2$ 1, $CaCl_2$ 1.8, glucose 10, hydroxyethyl piperazine ethane sulphonic acid (HEPES) 10 (pH 7.4). Single myocytes were whole-cell voltage-clamped using a pipette filled with intracellular-like solution containing (mM): NaCl 10, K-aspartate 130, $MgCl_2$ 1, EGTA 1, Na_2-ATP 2, HEPES 10 (pH 7.2). In experiments where adrenaline was used the pipette also contained 0.1 mM GTP, and ascorbic acid (0.5 mM) was added to the external solution. Pipettes for patch recording contained (mM): NaCl 70, KCl 70, $MgCl_2$ 1, $CaCl_2$ 1.8, $BaCl_2$ 1, $MnCl_2$ 2, HEPES 5 (pH 7.4). After a successful whole-cell recording of i_f was obtained in one cell, a second independent pipette was sealed onto the same cell for simultaneous cell-attached recording. Currents were recorded on tape and displayed on an 8-bit digital oscilloscope interfaced to a table computer. Single-channel currents were digitized at 500/1000 Hz and low-pass filtered before digital conversion using a 4-pole Bessel filter with a corner frequency of 80/160 Hz.

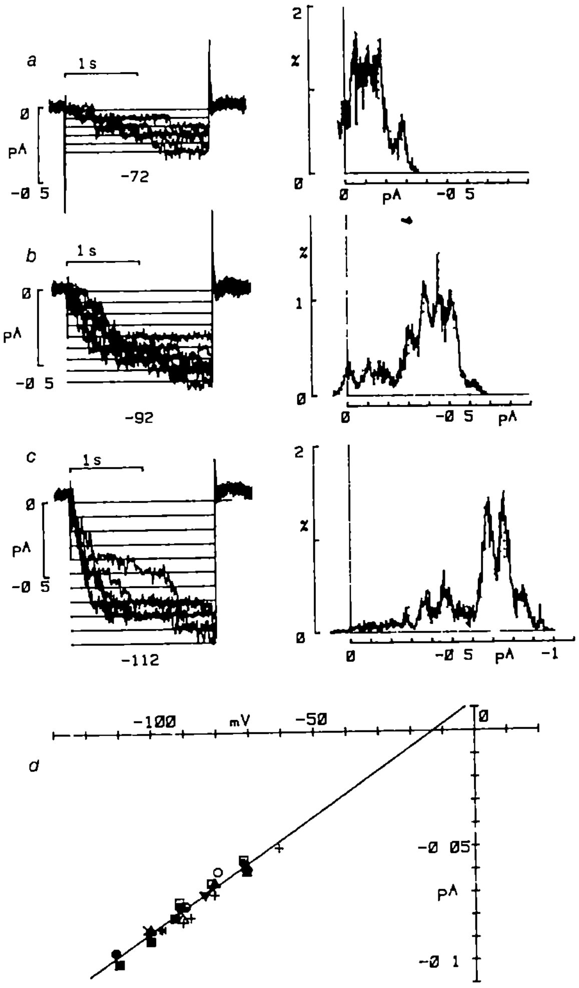


Fig. 2 *a–c*, transient distribution of patch currents during activation at different voltage levels; *d*, current-voltage relation for the single i_f channel. Whole-cell activation of i_f was obtained by applying 2 s hyperpolarizations to −72 (*a*), −92 (*b*) and −112 mV (*c*) from a holding potential of −32 mV through the pipette in whole-cell configuration. The left-hand panels show the superimposition of 4, 8 and 5 consecutive traces recorded from a membrane patch with a pipette in the cell-attached configuration sealed onto the same cell, during whole-cell hyperpolarizations. The right-hand panels show histograms of the current distribution for the same sets of patch records shown on the left. For each set of traces, the current axis was divided into intervals of 3 fA and the fraction of digitized points in each interval was plotted against current. Current levels corresponding to multiple channel openings can be clearly identified from the peaks in the histogram plots and correspond to single-channel levels evident in the current plots. Grids fitted by eye to histogram peaks (dotted lines) and current levels in the patch records (full lines) have interdistances of −0.056 pA (*a*), −0.075 pA (*b*) and −0.095 pA (*c*), representing the estimated single-i_f channel amplitudes at the absolute membrane voltages of −72, −92 and −112 mV, respectively. The maximum number of channels opening increased with larger hyperpolarizations, varying from 6 at −72 mV to 9 at −92 mV and 10 at −112 mV. This reflects the increase in the degree of i_f activation observed in whole-cell conditions[7,8]. *d*, Single-channel levels were measured in 10 cells by fitting either patch records or histograms as illustrated above in the voltage range −61 to −112 mV. Linear least-squares fitting of experimental points yields a single channel conductance of 0.98 pS and an extrapolated reversal at −13 mV.

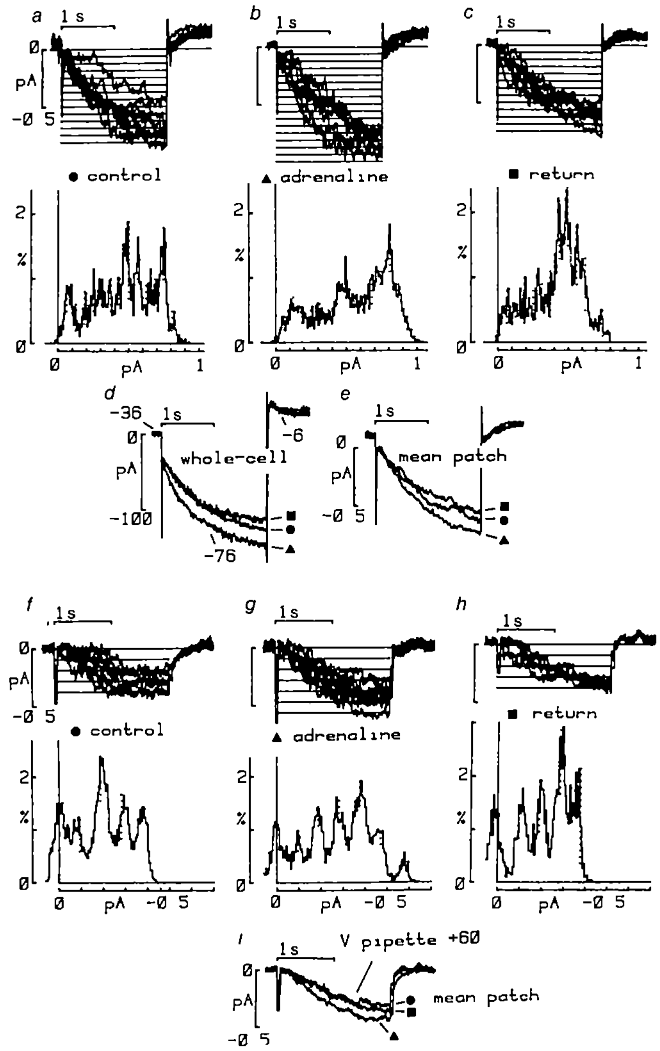


Fig 3 Modulation of i_f channels by adrenaline *a–e*, two-pipette experiment *f–i*, single-pipette experiment *a–e*, Whole-cell and patch currents were recorded during pulses to −76 mV from a holding potential of −36 mV in a control solution (*a*, ●), in a solution containing 5 μM adrenaline (*b*, ▲) and on return to the control solution after 2 5 min perfusion with the drug (*c*, ■) The single channel amplitude obtained from the fitting, done by eye, of patch current levels and histogram peaks did not vary in adrenaline (−0 062 pA) Adrenaline increased the maximal number of channels opened by hyperpolarizations (13 in *a*, 16 in *b* and 12 in *c*) and accelerated the rate of current development by decreasing the mean latency to first opening For example, the mean time required for 8 channels to be open was 452 ($n=7$), 326 ($n=12$) and 481 ms ($n=6$) in panels *a*, *b* and *c*, respectively This resulted in a faster activation rate of the mean patch current (*e*), which was comparable to that of the whole-cell current (*d*) Notice that the higher probability of channel opening favours the appearance of records with many open channels, and thus histogram peak levels corresponding to few open channels are depressed Histograms were constructed from 7 (*a*), 12 (*b*) and 6 (*c*) records *f–i*, Whole-cell pipettes were not used Patch currents were recorded during 2 s, +60 mV pulses applied to the cell-attached pipette Traces were electronically corrected for leakage and capacitative components and were stored after digital addition of the mean of 3 records during 2 s, −60 mV pulses Previous control runs showed that adrenaline at the concentration used did not alter significantly either the resting level or the holding current in the cell (see for example panel *d*) As in the two-pipette experiments, addition of 5 μM adrenaline increased the maximal number of channels opening (from 4 to 6) and caused i_f channels to open more readily (for example, the mean time required for the sequential opening of 3 channels was 670 ($n=6$), 521 ($n=9$) and 791 ms ($n=5$) in *f*, *g* and *h*, respectively) However, as apparent from the superimposition of patch traces and histograms, the elementary current step was unchanged in adrenaline (−0 094 pA) Histograms were constructed from 7 (*f*), 9 (*g*) and 5 (*h*) records

Recording of single-channel events of small amplitude during voltage pulses is complicated by the presence of capacitative transients and current drifts that cannot always be properly corrected To overcome this problem two pipettes have been used on the same cell, one for whole-cell and one for cell-attached patch recording Activation of i_f was obtained by applying intracellular hyperpolarizations via the whole-cell pipette as previously reported[19] Although technically more complicated, the use of two pipettes on one cell has several advantages over the single pipette patch-clamp method (1) voltage pulses delivered to the intracellular patch side do not involve charging of the cell-attached pipette, which eliminates the need for leakage and capacitative corrections and improves the resolution of single-channel current measurements, (2) single-channel activity and whole-cell currents are recorded simultaneously and can therefore be correlated, (3) the absolute membrane potential is known

A recent investigation[19] has shown that the current i_f in SA node cells has properties similar to those in the intact tissue The bottom traces in Fig 1 (labelled *d*) show whole-cell records of i_f activating during hyperpolarization to −82, −92 and −102 mV in one cell Simultaneous recording from a membrane patch of the same cell was performed using a second pipette in the cell-attached configuration, yielding records like those shown in traces *a* and *b* At all voltages the patch records show step-like transitions consistent with sequential openings of single inward channels that occur simultaneously with the whole-cell-driven activation of i_f

The elementary transitions that can be resolved in the cell-attached current records are of small amplitude −0 067 pA at −82 mV, −0 077 pA at −92 mV and −0 085 pA at −102 mV The activity recorded in the membrane patch reflects opening of single i_f channels, as shown by the similar time courses of averaged patch records (traces labelled *c*) and corresponding whole-cell i_f records (*d*) at the three voltages investigated

We studied the distribution of single open-channel current levels and determined the i_f elementary current step amplitudes during activation more accurately using the method shown in Fig 2 The current histogram for several cell-attached traces (left-hand panels) recorded during repetitive whole-cell-driven membrane hyperpolarizations was constructed at each voltage The histograms on the right give the fractional distributions of current levels for each set of traces At all voltages the histograms peak at the denser regions in the current panels, which represent the overlapping single-channel current levels Histogram peaks can be satisfactorily fitted by vertical grids with intervals of −0 056 pA at −72 mV, −0 075 pA at −92 mV and −0 095 pA at −112 mV In the lowest panel (*d*) the single-channel curve of current against voltage was reconstructed in the limited voltage range where measurements can be made using the experiments shown in Figs 1 and 2 and 8 further experiments The relationship is linear and the single channel conductance is found to be 0 98 pS The extrapolation crosses the voltage axis at −13 mV, a value close to that expected at the same external Na and K concentrations for i_f in Purkinje fibres[8]

Binding of catecholamines to cardiac β-receptors modulates pacemaker activity through the current i_f[3,10] Pacemaking acceleration is caused by a shift of the current activation curve to more positive voltages and is therefore attributable to an action of catecholamines on the i_f current kinetics[1,2,19,20] The effect of adrenaline on single i_f channels was investigated (Fig 3) The upper panels (a–e) show an experiment where adrenaline 5 μM was added to the cell superfusing solution In agreement with previous results[19], adrenaline caused the whole-cell i_f to increase during pulses to the current half-activation voltage range (panel d) A similar increase was also observed for the average of several patch records (e) Inspection of the patch records and corresponding histograms reveals that these changes are brought about by an increase of the maximal number of open channels (from 13 in a to 16 in b) and a decrease of the mean latency to first openings The single-channel amplitude was, on the other hand, unaltered (−0 062 pA throughout) These changes are in agreement with the shift in the current activation curve mentioned above

As shown in a recent study on i_f in SA node cells[19], part of the effectiveness of adrenaline on i_f may be lost due to buffering of cell content by the whole-cell pipette solution To avoid this, experiments were performed using cell-attached single-pipette patch-recording, thus excluding internal cell dialysis Cell-attached patch records obtained with the single-pipette arrangement needed to be corrected for leakage and capacitative components, and single i_f channels could be clearly resolved only in a few cases The lower panels in Fig 3 (f–i) illustrate one such case where in control conditions (f) four channels of −0 094 pA were seen to open during +60 mV pulses applied to the cell-attached pipette Perfusion with adrenaline (5 μM, g) caused more channels (6) to open more quickly compared with control conditions, without any alteration in the elementary current step size This resulted in a larger mean-patch current developing faster during a pulse (panel i), in agreement with the results of the two-pipette experiment above

The data presented here show that under favourable experimental conditions the opening of low-conductance single pacemaker channels during whole-cell driven activation of i_f can be resolved in isolated SA node cells In 70 mM KCl, 70 mM NaCl external solution the single i_f channel conductance is about 1 pS, a value much lower than that of other cation channels of low selectivity[21]

The recording of single i_f channels will be useful in detailed investigations of the properties of the pacemaker current, and in particular of the way the current is regulated by the adrenergic neurotransmitter Stimulation of cardiac β-adrenoceptors triggers a cascade of intracellular reactions leading to modifications of, among others, the slow Ca current and the pacemaker current[1,3,22–25], which are responsible for the chronotropic and inotropic responses of cardiac cells There is now ample evidence that modulation of cardiac Ca channels involves phosphorylation by the cAMP-dependent protein kinase[26–30] On the other hand, detailed information on the mechanisms involved in the catecholamine-controlled regulation of the current i_f has been limited so far by the absence of single i_f channel measurements The data presented here indicate that the pacemaker current modulation by adrenaline is mediated by an increased probability that i_f channels open upon hyperpolarization, while the channel conductance is unaltered These data confirm that adrenaline acts on the current kinetics rather than on its amplitude[1,2,20]

I thank Drs C Tromba and C Foglieni for assistance, Dr H F Brown for taking part in preliminary experiments and comments on the manuscript and Professor A Ferroni for encouragement This work was performed under a joint C N R and B C grant to D D and Professor D Noble (Oxford University) (CT85 00455 04 115 12810)

Received 12 May, accepted 8 October 1986

1 Hauswirth, O, Noble, D & Tsien, R W *Science* **162**, 916-917 (1968)
2 Tsien, R. W *J gen Physiol* **64**, 293-319 (1974)
3 Brown, H F, DiFrancesco, D & Noble, S J *Nature* **280**, 235-236 (1979)
4 Brown, H F & DiFrancesco, D *J Physiol, Lond* **308**, 331-351 (1980)
5 DiFrancesco, D & Ojeda, C *J Physiol, Lond* **308**, 353-367 (1980)
6 Yanagihara, K & Irisawa, H *Pflugers Arch. ges Physiol.* **385**, 11-19 (1980)
7 DiFrancesco, D *J Physiol, Lond* **314**, 359-376 (1981)
8 DiFrancesco D *J Physiol, Lond* **314**, 377-393 (1981)
9 Brown, H F *Physiol Rev* **62**, 505-530 (1982)
10 DiFrancesco, D *Prog Biophys molec. Biol.* **46**, 163-183 (1985)
11 Bader, C R & Bertrand, D *J Physiol, Lond.* **347**, 611-631 (1984)
12 Halliwell, J V & Adams, P R *Brain Res* **250**, 71-92 (1982)
13 Mayer, M L & Westbrook, G L *J Physiol, Lond.* **340**, 19-45 (1983)
14 Hamill, O P, Marty, A, Neher, E, Sakmann, B & Sigworth, F J *Pflugers Arch ges Physiol* **391**, 85-100 (1981)
15 Colquhoun, D, Neher, E, Reuter, H & Stevens, C F *Nature* **294**, 752-754 (1981)
16 Neher, E & Stevens, C F *A Rev Biophys Bioeng* **6**, 345-381 (1977)
17 Fenwick, E M, Marty, A & Neher, E *J Physiol, Lond* **331**, 577-597 (1982)
18 van Ginneken, A & Giles, W *Biophys J* **47**, 496a (1985)
19 DiFrancesco, D, Ferroni, A, Mazzanti, M & Tromba, C *J Physiol, Lond* **377**, 61-88 (1986)
20 Tsien, R W *J gen. Physiol.* **64**, 320-342 (1974)
21 Hille, B in *Ionic channels of excitable membranes*, 205-225 (Sinauer, Sunderland, 1984)
22 Tsien, R W, Giles, W & Greengard, P *Nature* **240**, 181-183 (1972)
23 Reuter, H *J Physiol., Lond.* **242**, 429-451 (1974)
24 Reuter, H & Sholtz, H *J Physiol., Lond.* **264**, 49-62 (1977)
25 Nargeot, J, Nerbonne, J M, Engels, J & Lester, H A *Proc. natn. Acad Sci USA* **80**, 2395-2399 (1983)
26 Ostrerneder, W *et al. Nature* **298**, 576-578 (1982)
27 Reuter, H *Nature* **301**, 569-574 (1983)
28 Brum, G, Flockerzi, V, Hofmann, F, Osterrieder, W & Trautwein, W *Pflügers Arch ges Physiol* **398**, 147-154 (1983)
29 Tsien, R. W *A Rev Physiol* **45**, 341-358 (1983)
30 Bean, B P, Nowycki, M C & Tsien, R W *Nature* **307**, 371-375 (1984)

Atrial natriuretic peptide causes pre-glomerular vasodilatation and post-glomerular vasoconstriction in rat kidney

M. Marin-Grez, J. T. Fleming & M. Steinhausen

Physiologisches Institut der Ludwigs-Maximilian Universitat,
Munchen and I Physiologisches Institut der Universitat,
Heidelberg, FRG

Atrial natriuretic peptide (ANP) can be extracted from rat hearts, and is found to increase fluid excretion by the kidneys when injected into test animals[1]. The mechanism of ANP action is still unclear. ANP may reduce sodium reabsorption in the renal tubules, but it is also known that it increases the rate of glomerular filtration in the kidney, and relaxes preparations of smooth muscle, including one made from arteries that supply the kidney[2,3]. To clarify its mode of action, we have studied directly the effects of semi-purified and synthetic ANP on blood vessels in the kidney of anaesthetized rats. We found that ANP causes a vasodilatation of the blood vessels which supply the glomeruli and a vasoconstriction of the arterioles which drain them. This substantiates the finding that increased filtration pressure participates in the natriuretic response[4].

The lack of effect of ANP on sodium transport *in vitro* suggests that reduction of sodium reabsorption *in vivo* may be secondary to another process, perhaps to renal vasodilatation Renal vasodilators cause natriuresis[5] ANP both lowers the vascular resistance of the pre-constricted isolated perfused kidney and increases renal blood flow (RBF) in intact animals[6,7] Suprarenal aortic clamping prevents the rise of GFR on ANP administration and blunts the natriuretic response, suggesting that vascular events are involved in this response[8]

However, RBF has also been found to be unaffected or even reduced after ANP administration[9,10] Since these conflicting effects could be due to different actions of ANP on distinct segments of the renal vascular tree, we investigated the effect of the peptide on pre- and post-glomerular resistance vessels

For this purpose semipurified ANP (see legend to Fig 2) or synthetic peptide (α-hANP) was infused intravenously into rats with unilateral hydronephrosis (Fig 1, refs 11 & 12) Experiments were also performed in which synthetic ANP (rANP, 26 aminoacids) was introduced into the lucite chamber holding the

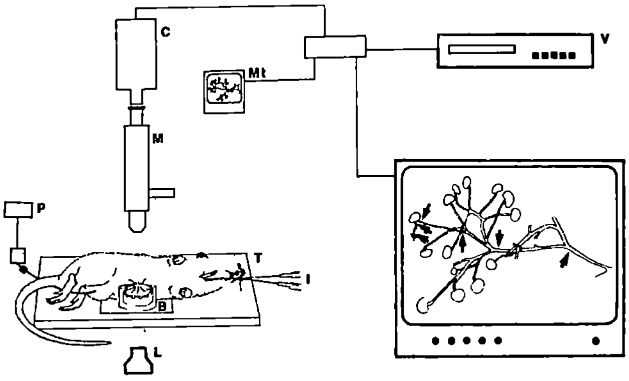


Fig. 1 Schematic representation of the experimental set-up L, light source, T, automatic controlled temperature table, B, temperature-controlled kidney bath, P, blood pressure transducer linked to amplifier and recorder, M, microscope, C, videocamera, V, video recorder, Mt, monitors The arrows on the monitor show the measuring points

Methods. For intravital microscopy a unilateral hydronephrosis was induced by occluding the renal artery (60 min) and ligating the ureter of female Wistar-Furth rats Seven to twelve weeks later the animals, which had been deprived of food but not of water for 15 h, were anesthetized with 5-ethyl-5-(1-methylpropyl)-2-thiobarbituric acid (Inactin, Byk Gulden, Konstanz, FRG, 100 mg kg^{-1}, i p). Body temperature was kept at 37 °C by a feed-back-controlled heated operating table The trachea was cannulated Blood pressure was recorded by means of a strain gauge connected to a catheter in the left carotid artery Infused solutions were delivered through two catheters into the left jugular vein The hydronephrotic kidney was exposed through a flank incision and split at its greater curvature using a cautery knife The dorsal sheet was retracted and the ventral one sutured to a semicircular wire holder This was fixed to a lucite chamber filled with warmed (37 °C) isotonic and isocolloidal solution (Haemaccel, Behringwerke, Marburg, FRG) The place at which the renal vessels entered the chamber was sealed with high vacuum grease (Dow Corning, Seneffe, Belgium) A vascular pathway that could be followed from the arcuate artery to the efferent arteriole was selected for study Measurements were made on the arcuate arteries (vessels which do not give rise to afferent arterioles) and on the interlobular arteries (vessels which do give rise to afferent arterioles), as well as on the afferent arteriole close to and distant from the glomerulus and on the efferent arteriole close to the glomerulus and at the welling point, by two independent observers Total magnification was 1389× and 2083×, respectively Isotonic NaCl was infused throughout the experiment at 3 75 ml h^{-1} An equilibration period of 2 h was allowed before starting the experiment

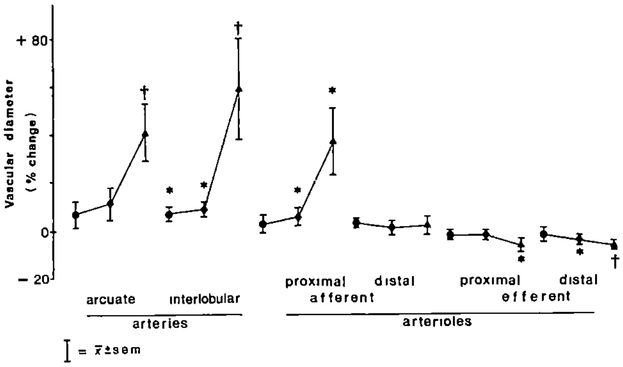


Fig 2 Percentage change of the vascular diameter at different levels of the renal vasculature Control measurements were made at the end of a 20 min i v infusion of potassium phosphate buffer (0 1 M, *p*H 7 4) at 10 or 100 μl/min^{-1} Semipurified ANP (see below) was infused for 20 min at the rate of ●, 1, ◆, 10 and ▲, 100 μl min^{-1} (Statistical significance is indicated thus *, $P<0\ 05$, †, $P<0\ 02$)

Methods. ANP was prepared from right and left atria dissected out of 44 rat hearts obtained under pentobarbital anesthesia (60 mg kg^{-1}, i p) The tissue was blotted, frozen in liquid nitrogen and pulverized Boiling water (32 ml) was added to the 3 2 g powder obtained After 15 min in a boiling water bath, the precipitate (11,000 g, 60 min, 4 °C) was discarded while the supernatant was removed and lyophilized The residue was dissolved in 20 ml phosphate buffer (0 67 M, *p*H 7 4), cold acetone was added (90%) and the precipitate was separated by centrifugation (11,000g, 30 min, 4 °C) The remaining acetone in the precipitate was evaporated at room temperature under vacuum. The residue was dissolved in 10 ml 150 mM NaCl and stored frozen (−20 °C) After thawing, the atrial extract was separated by chromatography in 2 separate runs through a BioGel P60 column (50 × 2 cm) Elution was carried out with potassium phosphate buffer (0 1 M, *p*H 7 4, 2 °C) at 10 ml h^{-1} The low molecular weight fractions containing natriuretic material were pooled An aliquot (0 1ml) of this pool increased sodium excretion by 69% in the rat bioassay[21] Considering that atrial ANP content is 10–20 μg[22] and that about 50% natriuretic activity is recovered in the low molecular weight fractions, our preparation should have contained about 100–200 ng ANP ml^{-1} so that bolus injection of 0 1 ml should increase sodium excretion by 50–100%[23] The observed value was within this range (see above)

kidney Images of intrarenal vessels were formed in a television camera, through a microscope immersion objective, and displayed on monitors for on-line measurements of the lumen diameter Each experiment was recorded on video tape for off-line analysis The diameter of a given vessel, measured during an infusion of ANP or after application of the peptide to the lucite chamber holding the kidney, was expressed as a percentage of the diameter measured during the control period The data was assessed by analysis of variance and the statistical significance evaluated by Wilcoxon's test for paired differences Results are given as mean ± s e m

A dose-dependent vasodilatation of the arcuate and interlobular arteries was observed during the infusion of semipurified ANP (Fig 2, Table 1) The vascular diameter was 141% and 159% of control with the highest dose respectively, indicating a decrease in vascular resistance to less than 30% and 20% of control respectively The proximal end of the afferent arteriole was also vasodilated by ANP (Fig 2) The effect (137% of control with the highest dose infused) was, though significant, somewhat lower than that seen in the interlobular artery Semipurified low molecular weight ANP had no effect on the afferent arteriole in the vicinity of the glomerulus The efferent arteriole was constricted by the high dose near the glomerulus and by the middle and high dose away from it The estimated increases in vascular resistance with the high dose of semipurified ANP was 35% near the glomerulus and 64% distant from it (Fig 2) Blood pressure decreased from 112 1 ± 3 0 mm Hg during the control period to 102 1 ± 3 2 mm Hg ($P<0\ 01$) during the 100 μl min^{-1} ANP infusion It remained unchanged (113 ± 5 and 109 ± 4 mm Hg) at 1 and 10 μl min^{-1} ANP

Infusion of synthetic α-hANP dilated all preglomerular vessels and constricted the efferent arteriole (Fig 3) The beginning of the afferent arteriole was the segment most strongly vasodilated, its diameter increasing by 12 ± 8%, 38 ± 13% and 42 ± 6% at infusion rates of 0 05, 0 1 and 0 2 μg α-hANP min^{-1} respectively ($n=3$) The distal portion of the efferent arteriole was vasoconstricted by 5 ± 0 3% and 5 9 ± 1% with 0 1 and 0 2 μg α-hANP min^{-1} respectively The blood pressure fell from 95 mm Hg during the control phase to 90 ± 4 mm Hg during the infusion of 0 2 μg min^{-1} ANP Addition of synthetic rANP to the lucite bath also induced a pre-glomerular vasodilatation and a post-glomerular vasoconstriction (Table 2) without affecting the arterial blood pressure

Our results clearly show that ANP, of either natural or synthetic origin, has the capacity to induce vasodilatation in the preglomerular renal vessels Our findings are in agreement with

Table 1 Diameter of intrarenal vessels (μm) during infusion of semipurified ANP

	Control phase	Treatment rate of infusion of ANP (μl ANP pool min^{-1}) 1	10	100
Arcuate artery	25.0±1.6	26.6±1.6	27.6±1.9	32.4±2.1
Interlobular artery	11.7±1.2	12.5±1.3	12.7±1.3	18.1±1.3
Afferent arteriole, proximal end	6.6±0.9	6.6±0.8	6.9±0.8	9.5±1.0
Afferent arteriole, distal end	6.0±0.6	6.2±0.6	6.1±0.6	6.2±0.7
Efferent arteriole, proximal end	12.4±1.0	12.1±0.9	12.4±1.1	11.7±1.0
Efferent arteriole, distal end	16.9±1.7	16.3±1.4	16.1±1.5	15.2±1.4

Semipurified ANP was prepared as described for Fig. 2. Each value is based on 7 measurements.

Table 2 Percentage change in vascular diameter in the hydronephrotic kidney after treatment with synthetic ANP

Location of measurement	Concentration of rANP in chamber (mol l^{-1}) 10^{-9}	10^{-8}	10^{-7}	3×10^{-7}
Interlobular artery, near arcuate artery	+9.7±2.9* (6)	±17.0±7.4 (6)	+25.5±7.3* (6)	+39.0±11.9* (6)
Interlobular artery, near afferent arteriole	+3.6±3.6 (6)	+12.5±3.4* (6)	+27.6±6.8* (6)	+31.6±7.5* (6)
Afferent arteriole, near interlobular artery	+12.0±3.6* (6)	+18.0±3.8† (6)	+32.8±7.7* (6)	+30.0±4.6† (6)
Afferent arteriole, near glomerulus	+10.5±2.6* (6)	+14.2±2.9† (6)	+20.0±3.9† (6)	+28.1±7.9* (6)
Efferent arteriole, near glomerulus	2.9±2.0 (12)	−5.5±3.0 (12)	−10.0±3.0* (18)	
Efferent arteriole, near welling point	−2.0±3.0 (12)	−6.0±5.0 (12)	−12.0±3.0* (18)	

Statistical significance as follows: * $P < 0.05$, † $P < 0.01$. Measurements were made 15 min after adding rANP (R-R-S-S-C-F-G-G-R-I-D-R-I-G-A-Q-S-G-L-G-C-N-S-F-R-Y) to the lucite chamber holding the kidney. Figures in parenthesis indicate the number of measurements used to obtain each value.

the observation that ANP reduces the induced tone of isolated renal arcuate arteries from rat (2). We cannot exclude the possibility that in the hydronephrotic kidney the preglomerular vessels are partly vasoconstricted and thus more sensitive to vasodilatory stimuli. However, angiotensin (0.2 μg per kg per min) still constricts preglomerular vessels by 30–40% in this system[13]. It has been suggested that the natriuretic effect of ANP depends mainly on enhanced GFR[14]. We have found in addition that this peptide increases proximal tubule stop-flow pressure (an index of filtration pressure, ref. 4). It appears therefore reasonable to conclude that in intact animals ANP also causes pre-glomerular vasodilatation. ANP may, in addition, induce post-glomerular vasoconstriction, and even the slight reduction in diameter seen with the lowest dose of ANP tested can cause a substantial rise in resistance. It has been shown that in the absence of vasoconstrictors ANP increases the renal vascular resistance and the glomerular filtration rate of the isolated perfused rat kidney, suggesting a preferential effect on the efferent arterioles[15]. Increased filtration fraction has been a repeated finding in studies of the renal effects of atrial peptides[16].

The vasodilating effect of ANP might be due to activation of the particulate guanylate cyclase[17]. Involvement of dopamine cannot yet be excluded. Although dopamine antagonists like haloperidol block the natriuretic effect of ANP, they do not prevent the vasodilating effect of the peptide[18,19].

It seems likely that the primary role of ANP is to regulate renal haemodynamics, and that ANP-induced natriuresis depends on increased filtration pressure and dissipation of the medullary osmotic gradient[10].

This work was supported by the Deutsche Forschungsgemeinschaft Forschergruppe Niere. Dr Fleming was a fellow of the Boehringer Ingelheim Fonds. Synthetic rANP was a gift of Merck-Sharp & Dohme. We acknowledge the excellent technical assistance of R. Dussel and E. Mack-Kühn.

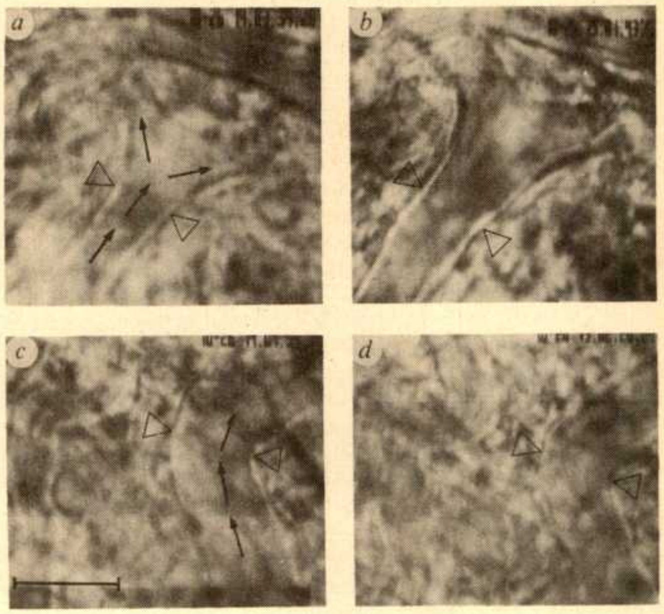


Fig. 3 Intravital microphotographs of split hydronephrotic kidney. *a*, Interlobular artery at the branching point into two afferent arterioles; *b*, same vessel after administration of 158 ng min^{-1} α-hANP; *c*, efferent arteriole just before the branching point, *d*, same vessel after the administration of 158 ng min^{-1} α-hANP. The flow direction is marked by arrows. Arrowheads, the same site of the vessel. Scale bar, 30 μm.

Received 29 April; accepted 6 October 1986.

1. De Bold, A. J., Borstein, H. B., Veress, A. T. & Sonnenberg, H. *Life Sci.* **28,** 89-94 (1981).
2. Aalkjaer, C., Mulvany, M. J. & Nyborg, N. C. B. *Br. J. Pharmac.* **86,** 447-453 (1985).
3. Huang, C.-L., Lewicki, J., Johnson, L. K. & Cogan, M. G. *J. clin. Invest.* **75,** 769-773 (1985).
4. Schnermann, J., Marin-Grez, M. & Briggs, J. P. *Pflügers Arch. ges. Physiol.* **406,** 237-239 (1986).
5. Barrachlough, M. A. & Mills, I. H. *Clin. Sci.* **28,** 69-74 (1965).
6. Baines, A. D., de Bold, A. J. & Sonnenberg, H. *Can. J. Physiol. Pharmac.* **61,** 1462-1466 (1983).
7. Wakitani, K. *et al. Am. J. Physiol.* **249,** F49-F53 (1985).
8. Burnett, J. C., Jr, Opgenorth, T. J. & Granger, J. P. *Kidney Int.* **30,** 16-19 (1986).
9. Camargo, M. J. F., Kleinert, H. D., Atlas, S. A., Sealey, J. E. & Laragh, J. H. *Am. J. Physiol.* **246,** F447-F456 (1984).
10. Seymour, A. A., Smith, S. G., Mazack, E. K. & Blaine, E. H. *Hypertension* **8,** 211-216 (1986).
11. Steinhausen, M., Snoei, H., Parekh, N., Baker, R. & Johnson, P. C. *Kidney Int.* **23,** 794-806 (1983).
12. Steinhausen, M. *et al. Kidney Int.* **30,** 56-61 (1986); *Kidney Int.* **30,** 361-370 (1986).
13. Steinhausen, M. *et al. Kidney Int.* (in the press).
14. Steigerwall, S., Carretero, O. A. & Beierwaltes, W. H. *Proc. Soc. exp. Biol. Med.* **182,** 88-94 (1986).
15. Atlas, S. A. *et al. Nature* **309,** 717-719 (1984).
16. Needleman, P. *et al. Hypertension* **7,** 469-482 (1985).
17. Waldman, S. A., Rapoport, R. M. & Murad, F. *J. biol. Chem.* **259,** 14332-14334 (1984).
18. Marin-Grez, M., Briggs, J. P., Schubert, G. & Schnermann, J. *Life Sci.* **36,** 2171-2176 (1985).
19. Oschima, T., Currie, M. G., Geller, D. M. & Needleman, P. *Circulation Res.* **54,** 612-616 (1984).
20. Maack, T. *et al. Am. J. Med.* **61,** 1069-1075 (1984).
21. Briggs, J. P., Marin-Grez, M., Steipe, B., Schubert, G. & Schnermann, J. *Am. J. Physiol.* **247,** F480-F484 (1984).
22. Lang, R. E. *et al. Nature* **314,** 264-266 (1985).
23. Sugiyama, N. *et al. Biochem. biophys. Res. Commun.* **123,** 338-344 (1984).

Pancreastatin, a novel pancreatic peptide that inhibits insulin secretion

Kazuhiko Tatemoto*, Saud Efendić†, Viktor Mutt‡, George Makk*, Gottfried J. Feistner* & Jack D. Barchas*

* Nancy Pritzker Laboratory, Department of Psychiatry and
Behavioral Sciences, Stanford University School of Medicine,
Stanford, California 94305, USA
† Department of Endocrinology, Karolinska Hospital and
‡ Department of Biochemistry II, Karolinska Institute,
S-104 01 Stockholm, Sweden

In mammalian tissues the C-terminal amide structure has been found to occur only in neuroactive or hormonally-active peptides[1]. About half known neuropeptide and peptide hormones have this unique chemical feature. Using a chemical detection method[2], a search for previously unknown peptides that possess the C-terminal amide structure in extracts of brain and intestine was carried out[3] and a number of novel neuropeptides and hormonal peptides[4], designated neuropeptide Y[5,6], PHI[7], peptide YY[8], galanin[9] and neuropeptide K[10] were isolated. We recently performed a similar search in porcine pancreas and found a high concentration of a peptide having a glycine amide at its C-terminus. Here we report the isolation, primary structure and biological activity of this novel peptide. The 49-residue peptide strongly inhibits glucose-induced insulin release from the isolated perfused pancreas and was therefore named pancreastatin. It may be important in the regulation of insulin secretion and in the pathogenesis and treatment of diabetes mellitus.

The pancreastatin-containing fractions were located by the amounts of glycine amide released from tryptic digests of the samples using the chemical method described previously[2]. The peptide was extracted from porcine pancreas (20 kg) as described[11], and purified by ethanol fractionation, gel filtration (Sephadex G-25) and ion exchange chromatography (CM cellulose) (Fig. 1, legend). The final HPLC purification steps (Fig. 1*a*-*c*) yielded a homogeneous peptide preparation (0.2 mg) which was subjected to structural analysis.

Terminal analyses indicated that the peptide has an N-terminal glycine and a C-terminal glycine amide. From amino acid analysis, it consists of 49 residues: Ala (8), Arg (3), Asp (1), Glu (13), Gly (9), His (1), Leu (1), Lys (2), Met (1), Phe (1), Pro (5), Ser (1), Thr (2), Trp (1). Edman degradation of the intact molecule (2 nmoles) using a gas-phase sequencer (Applied Biosystems) revealed the sequence of the first 25 residues of the N-terminal region. Because there are only two lysine residues in the molecule, the intact peptide (1.8 nmoles) was treated with a lysine-specific protease (Endoproteinase Lys-C, Boehringer, Mannheim) and the resulting digest was separated on HPLC (Fig. 1*d*). Peak I was found to contain a peptide with an N-terminal threonine. Sequence analysis indicated that this peptide corresponded to the middle fragment (residues 14-25) of the parent molecule. Peak II was found to contain two peptides, one with an N-terminal glycine and the other with an N-terminal glycine and a C-terminal glycine amide, indicating that this peak contained both the N- and C-terminal fragments of the parent molecule. Edman degradation of the peptides in peak II therefore yielded two residues for each cycle. Because the structure of the N-terminal fragment (residues 1-13) was already known, the structure of residues 26-49 was deduced from the sequence data by subtraction of the N-terminal sequence. The complete amino-acid sequence of porcine pancreastatin was deduced in this way (Fig. 2). In agreement with the proposed structure, pancreastatin was found to have a chemical molecular weight of 5103.1 daltons (calculated 5103.46) as measured by FAB mass-spectrometry (VG-ZAB-SE, R700, 10 μg pancreastatin in thioglycerol, mass range 5120-4780 u, scan rate 200 s/d, 15 scans accumulated).

Table 1 Effect of pancreastatin and its fragments (10 mM) on glucose-induced insulin response of the isolated rat pancreas

Conditions of perfusion	*n*	Insulin release (μU min^{-1} ml^{-1})	
		1-5 min	5-20 min
control	5	1,764 ± 210	227,184 ± 2,515
pancreastatin	6	676 ± 74‡	21,134 ± 2,288
fragment 33-49	5	817 ± 178†	15,423 ± 3,524*
fragment 14-49	5	627 ± 116†	17,056 ± 3,114*

† $P < 0.01$, ‡ $P < 0.001$ v.s. contols.
* $P < 0.05$.

The primary structure of pancreastatin is different from all other known peptides although the peptide has partial structural similarities to several known peptides. For example, pancreastatin shares the -Glu-Glu-Glu-Glu-Glu-structure (residues 34-38) with gastrin[12], and the C-terminal -Arg-Gly-NH_2 structure with vasopressin[13]. However, no other structural homology with these peptides is evident. These results suggest that pancreastatin may belong to a hitherto unknown peptide family.

Chemical synthesis of pancreastatin and its fragments was carried out using a solid-phase synthesizer. Fully protected pancreastatin (1-49) and the (14-49) and (33-49) fragments were synthesized by a published method[14]. After complete deprotection and cleavage by hydrofluoric acid, the crude synthetic preparations were purified by semipreparative HPLC. The synthetic pancreastatin (1-49) obtained was found to elute from the HPLC at the same position as the native peptide (Fig. 1). The results of amino-acid sequence analysis and molecular weight (found 5103.4) determination by FAB mass-spectrometry indicated that the synthetic peptide had an identical structure to native pancreastatin.

The effects of natural and synthetic pancreastatin and the synthetic fragments, pancreastatin (14-49) and pancreastatin (33-49) on glucose-stimulated insulin release, have been studied using isolated rat pancreas[15]. As expected, perfusion with glucose (16.7 mM) induced a biphasic insulin release from the isolated pancreas. Pancreastatin and the two fragments (10 nM) markedly decreased the early phase of insulin release (Fig. 3 and Table 1). The effect on late-phase insulin secretion was less

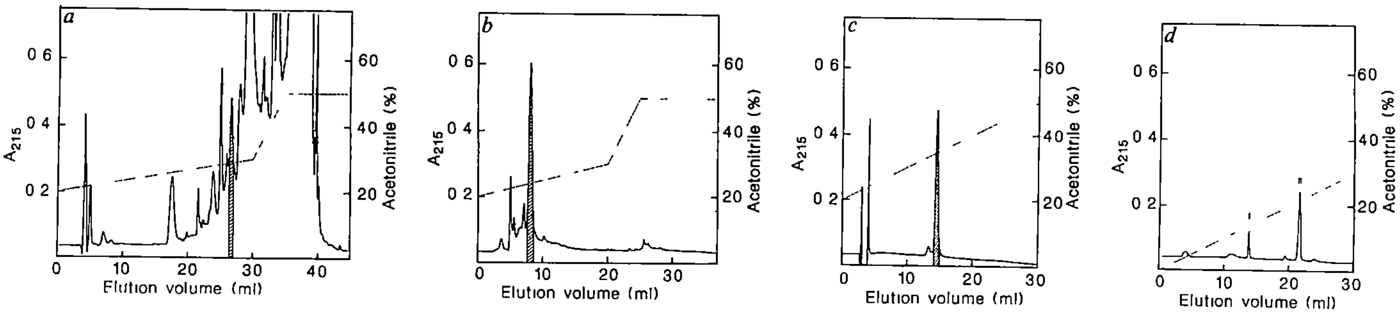


Fig. 1 Reverse-phase HPLC purification of pancreastatin (*a–c*) and its fragments (*d*). *a*, Reverse-phase HPLC of the partially-purified pancreastatin preparation (5 mg, see below) on a Ultrapac TSK ODS 120-T column (4.6×250 mm, LKB AB, Sweden) using a gradient of acetonitrile in 0.1% aqueous trifluoroacetic acid at a flow rate of 1.0 ml min^{-1}. Absorbance at 215 mm (A_{215}) is shown. The pancreastatin-containing fraction (hatched area) was evaporated to dryness using a vacuum centrifuge. A total of 20 mg of the preparation was purified by repeating the HPLC separation and the combined pancreastatin fractions were subjected to the next HPLC purification. *b*, Reverse-phase HPLC of the pancreastatin fractions from *a* on a Ultrapac TSK ODS 120-T column with a gradient of acetonitrile in 10 mM phosphate buffer *p*H 6.5 (flow rate 1.0 ml min^{-1}). *c*, Reverse-phase HPLC of the pancreastatin fraction from *b* on a Ultrasphere ODS column (4.6×250 mm, Altex) under the conditions described in *a*. *d*, HPLC separation of the peptide fragments obtained by treatment of pancreastatin with a lysine-specific protease. Pancreastatin (1.8 nmoles) was dissolved in 10 μl of 1% ammonium bicarbonate, 2 μl of Endoproteinase Lys-C (Boehringer Mannheim) solution (2 mg ml^{-1}) was added and the reaction mixture was incubated at room temperature overnight, then at 100 °C for 10 min and lyophilized. The enzyme digest was applied to a reverse-phase HPLC column (Ultrapac TSK-ODS 120-T, 4.6×250 mm) and eluted at a flow rate of 1 ml min^{-1} with the solvent systems described in *a*. The partially-purified pancreastatin preparation was prepared by boiling porcine pancreas (20 kg) in water for 10 min. The pancreas was then frozen, minced and extracted with 0.5 M acetic acid for 24 h. Peptides in the extract were adsorbed onto alginic acid (1 kg wet weight (wt)), eluted with 0.2 M HCl and precipitated with NaCl at saturation. The peptide precipitate (18 g wet wt) was redissolved in water (360 ml) and two volumes of ethanol were added to the solution. The *p*H of the solution was adjusted to 7.2 and the resulting precipitate was removed by filtration. After addition of three volumes of water to the filtrate, peptides in the filtrate were again adsorbed onto alginic acid (250 g wet wt), eluted with 0.2 M HCl and reprecipitated with NaCl at saturation. This second peptide precipitate (4 g wet wt) was applied to a Sephadex G-25 column and eluted with 0.2 M acetic acid. Fractions containing pancreastatin were pooled and lyophilized. The lyophilized material (0.8 g) was further purified by ion-exchange chromatography on CM cellulose with a step-wise elution (0.01–0.2 M) of ammonium bicarbonate. The fraction eluted by 0.01 M ammonium bicarbonate was found to contain pancreastatin and this partially-purified preparation (90 mg) was used as the starting material for further purification by the HPLC steps.

1 5 10 15
Gly-Trp-Pro-Gln-Ala-Pro-Ala-Met-Asp-Gly-Ala-Gly-Lys-Thr-Gly-Ala-Glu-
↑
20 25 30
Glu-Ala-Gln-Pro-Pro-Glu-Gly-Lys-Gly-Ala-Arg-Glu-His-Ser-Arg-Gln-Glu-
↑
35 40 45
Glu-Glu-Glu-Glu-Thr-Ala-Gly-Ala-Pro-Gln-Gly-Leu-Phe-Arg-Gly-NH$_2$

Fig. 2 Complete amino-acid sequence of pancreastatin from porcine pancreas. The cleavage sites for Endoproteinase Lys-C are indicated by the arrows.

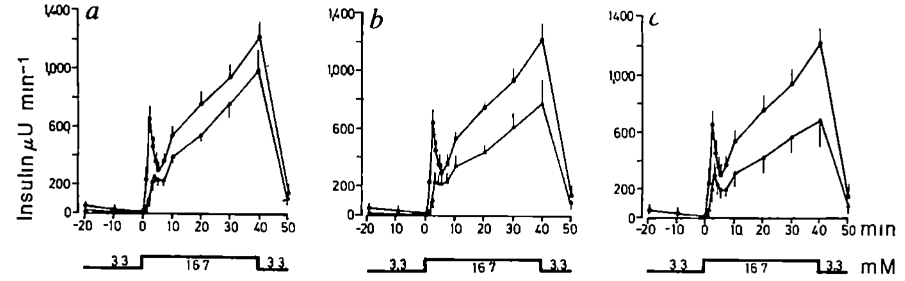


Fig. 3 The effect of 10 nM of *a*, pancreastatin (1–49), *b*, pancreastatin (14–49) and *c*, pancreastatin (33–49) on the glucose-induced insulin release from the isolated perfused rat pancreas. ●, control values, ▲, test values. The insulin release in the presence of pancreastatin or fragment is lower in each case than the control. The perfused rat pancreas was prepared as follows. Sprague-Dawley rats, weighing 200–250 g, were fed *ad lib*. They were anaesthetized by intraperitoneal injection of 50 mg kg^{-1} pentobarbital and the pancreas was isolated[15]. In control experiments the pancreas was first perfused with Krebs-Ringer buffer containing 20 g l^{-1} of bovine serum albumin and 3.3 mM glucose for 20 min and then perfused for 50 min with the buffer and 16.7 mM glucose. The peptide (10 mM) was added 10 min before and during glucose (16.7 mM) administration. The solution was perfused by the use of a perfusion pump through the abdominal aorta. The flow rate of perfusion was 3 ml min^{-1}. The perfusate was collected in tubes containing 0.1 ml of Trasylol. Insulin was measured by radioimmunoassay using a rat insulin standard. The sensitivity of the assay was about 300 ng l^{-1} and coefficient of variation 10%. Points are the mean ± s.e.m. for 5 determinations, except for the test valves in *a*, which are the mean of 6 determinations.

pronounced, but statistically significant for the fragments Thus pancreastatin may have an important role in the regulation of insulin secretion The potent effects of the C-terminal fragments of pancreastatin on insulin release suggest that this portion of the molecule is important for the biological activity and that C-terminal active fragments of pancreastatin may occur *in vivo* It has been suggested[16] that suppression of insulin release upon glucose stimulation, especially in the early phase, is a characteristic feature of manifest type-2 diabetes Because pancreastatin strongly suppresses insulin release, particularly in the early phase, it may be important to examine whether abnormalities in the regulation or action of pancreastatin and its receptors are involved in the pathogenesis of type-2 diabetes If so, the discovery of pancreastatin would have therapeutic implications, particularly in the treatment of diabetes mellitus

Preliminary results indicate that pancreastatin also inhibits the release of somatostatin upon glucose stimulation from the perfused pancreas Pancreastatin may thus also be important in the control of carbohydrate metabolism and hyperglycaemia in diabetes

We thank Drs S R Levin and K Faull for discussions and Mrs P Angwin, A Nylen, L Melin and Mr W Kammerer for technical assistance This study was supported by grants from NIMH (MH 23861), the Office of Naval Research, the Swedish Medical Research Council (Projects 00034 and 13X-1010) and Nordisk Insulinfond

Received 7 July, accepted 3 October 1986

1 Narita, K, Matsuo, H & Nakajima, T in *Protein Sequence Determination* (ed Needleman, S B) 30–130 (Springer, Berlin, 1975)
2 Tatemoto, K & Mutt, V *Proc natn Acad Sci USA* **75**, 4115-4119 (1978)
3 Tatemoto, K & Mutt, V *Nature* **285**, 417-418 (1980)
4 Tatemoto, K *Biomed. Res* **4**, Suppl 1, 1-6 (1983)
5 Tatemoto, K, Carlquist, M & Mutt, V *Nature* **296**, 659-660 (1982)
6 Tatemoto, K *Proc natn Acad Sci USA* **79**, 5485-5489 (1982)
7 Tatemoto, K & Mutt, V *Proc natn Acad Sci USA* **78**, 6603-6607 (1981)
8 Tatemoto, K *Proc natn Acad Sci USA* **79**, 2514-2518 (1982)
9 Tatemoto, K, Rokaeus, A, Jornvall, H, McDonald, T J & Mutt, V *FEBS Lett* **164**, 124-128 (1983)
10 Tatemoto, K, Lundberg, J M, Jornvall, H & Mutt, V *Biochem. biophys Res Commun* **128**, 947-953 (1985)
11 Jorpes, J E & Mutt, V US patent 3,013,944 (1961)
12 Gregory, H, Hardy, P M, Jones, D S, Kenner, G W & Sheppard, R C *Nature* **204**, 931-933 (1964)
13 du Vigneaud, V, Lawler, H C & Popenoe, A *J Am chem. Soc* 4880-4881 (1953)
14 Merrifield R B *J Am chem Soc* **85**, 2149-2154 (1963)
15 Loubatieres, A, Mariani, M M, Ribes, G, de Malbose, H & Chapel, J *Diabetologia* **5**, 1-10 (1969)
16 Efendic, S, Luft, R & Wajngot, A *Endocrine Rev* **5**, 395-410 (1984)

Glucocorticoid receptors lacking hormone-binding activity are bound in nuclei of ATP-depleted cells

Dirk B. Mendel, Jack E. Bodwell & Allan Munck

Department of Physiology, Dartmouth Medical School,
Hanover, New Hampshire 03756, USA

The glucocorticoid receptor binding capacity of rat thymus cells disappears when the cells are depleted of ATP by anaerobiosis, and rapidly reappears when ATP levels are restored[1–3]. Loss and recovery of binding capacity occurs even when protein synthesis is suppressed with cycloheximide[3,4]. In view of this and similar work in other cell systems[5,6], we proposed that in cells deprived of ATP the receptor is present in a form—the 'null receptor' form, as we shall call it—that cannot bind hormone[4]. Although many subsequent observations support this idea, no direct evidence has appeared for the existence of the null receptor. We have attempted to detect the null receptor in WEHI-7 mouse thymoma cells with a monoclonal antibody to the glucocorticoid receptor. Here we report that the null receptor is bound in the nuclei of ATP-depleted cells, and is present in amounts comparable to those of receptors in normal cells

WEHI-7 cells incubated without hormone at 37 °C with 1 mM dinitrophenol (DNP) rapidly lose ATP and receptor binding capacity (Fig 1) ATP levels drop to ~10% of control values within 10 min, whereas glucocorticoid receptor binding capacity in the cell cytosol decreases more slowly No significant decreases occur in control cells incubated without DNP These results are consistent with those obtained with other cells and other methods for lowering ATP[1-6] Receptor binding measured with whole cells falls similarly[1-6], and, as we show below, no binding is detectable in nuclei from the DNP-treated cells

Although there is controversy over the intracellular location of glucocorticoid receptors in the absence of steroid[7-9], by most cell fractionation procedures (including ours) the receptors without bound ligand are recovered in the cytosolic fraction Because our experiments were conducted without hormone, we first looked for the null receptor in that fraction Our attempts to immunoprecipitate receptors from cytosolic fractions from cells incubated with DNP for 90 min were unsuccessful, leading us to suspect that the null receptor, if it existed at all, was located in the nuclear fraction DNP-treated WEHI-7 cells were lysed using a freeze-thaw procedure by which receptors from cytosolic and nuclear fractions can be purified separately The relative amounts of receptor in the two fractions were determined from the relative amounts of immunoreactive protein detected by Western blot analysis[10]

Lanes 1 and 2 of Fig 2 represent cytosolic and nuclear fractions from cells lysed immediately after the addition of DNP Almost all the receptors (arrow) are located in the cytosolic fraction The small amount of receptor in the nuclear fraction (1,800 c p m ^{125}I in lane 2 compared with 15,000 c p m in lane 1) may represent trace contamination with cytosolic receptor This result is similar to that obtained with WEHI-7 cells in the absence of steroid and DNP However, after 90 min of DNP treatment (Fig 2, lanes 3 and 4), the receptors are located predominantly in the nuclear fraction (lane 4), even without hormone present The cytosolic fraction (lane 3), which contained ~20% of the original binding capacity (Fig 1*b*), contained a similar proportion of the immunoreactive receptor protein (5,000 c p m in lane 3 compared with 21,000 c p m in lane 4) The total amounts of receptor present in the cells before and after DNP treatment were comparable (17,000 c p m in lanes 1 and 2 and 26,000 c p m in lanes 3 and 4, the greater amount of receptor after DNP treatment probably indicates more efficient recovery of receptor from the nuclei than the cytosol) The receptors in the DNP-treated cells do not appear to be bound to the outer nuclear membranes as they were still present in nuclei treated with 0 05% Triton X-100 for 5 min at 3 °C to strip off the outer nuclear membranes[11] (23,000 c p m in lane 5, 21,000 c p m in lane 4) Preliminary studies, in fact, suggest that the null receptor is less easily extracted from nuclei than most nuclear-bound complexes formed by glucocorticoids

In separate experiments (not shown), nuclei of DNP-treated cells were incubated for 2 h at 3 °C with 50 nM ^{3}H-triamcinolone acetonide, in the presence and absence of 2 μM of the unlabelled compound, to test the ability of these receptors to bind hormone No saturable binding was detected, indicating that the null receptor cannot bind steroid.

Nuclear binding of steroid hormone receptors in whole cells has until now been induced only with hormone, which forms receptor complexes that become activated and bind to the nucleus New light may be shed on these processes by our observation that nuclear binding can be induced without hormone Initially we proposed that the null receptor represents a stage of a cycle, possibly driven by cyclic phosphorylation-dephosphorylation[4], in which the receptor, after binding in the

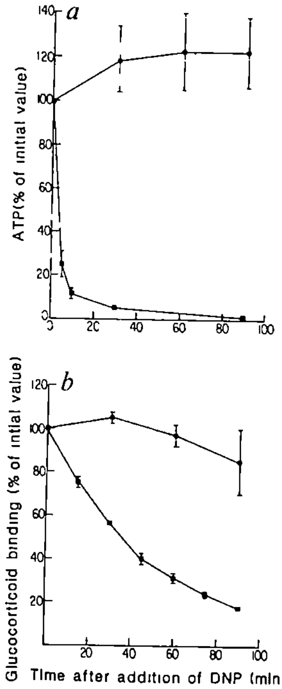


Fig 1 Time course of drop in cellular ATP and glucocorticoid-binding capacity in WEHI-7 cells WEHI-7 cells ($\sim 2\times10^8$), maintained in culture as described[10], were washed by suspending in 80 ml of modified Krebs-Ringer buffer containing 25 mM hydroxyethyl piperazine ethane sulphonic acid (HEPES), *p*H 7 4, 3 6 mM sodium bicarbonate, 2 mg ml^{-1} bovine serum albumin, at 37 °C for 5 min, the suspension centrifuged at 400*g* for 3 min, and the cells resuspended in 3 6 ml of the same buffer containing 1 mM DNP (■) or no DNP (●) Cells were incubated at 37 °C in an orbital water bath At various times, aliquots of the cell suspension were assayed for ATP (*a*) and saturable glucocorticoid-binding capacity (*b*) Results are presented as % of initial value determined at 0 min, just after the addition of DNP, and represent the main value ± s e m of duplicate measurements from two separate experiments Cell viability, as determined by Trypan blue exclusion, was unaffected by the 90 min DNP treatment

Methods. *a*, Cellular ATP was assayed by a modification of the luciferase procedure[16] 25 μl of the cell suspension was added to 500 μl of 5% perchloric acid at 0 °C After 30 min, the sample was centrifuged at 11,000 *g* for 4 min Aliquots (450 μl) of the supernatant were added to 500 μl of a 1 1 mixture of trichlorotrifluoroethane and tri-*n*-octylamine[17] The mixture was vortexed and centriguged at 1,500*g* for 2 min in a clinical centrifuge at room temperature The amount of ATP in 20 μl aliquots of the top phase, which contained the nucleotides in acid-free solution, was determined in a Packard liquid scintillation spectrophotometer from the light emitted upon addition of 100 μl of a solution of desiccated firefly tails (Sigma) prepared in sodium arsenate according to the manufacturer's instructions Absolute values of ATP were determined from a standard curve of events per 0 1 min versus known quantities of ATP prepared from a 2 mM stock solution *b*, Glucocorticoid receptor binding capacity was assayed in cytosolic fractions obtained by a modification of the freeze-thaw procedure[10] in which 100 μl of the cell suspension was added to 250 μl of lysis buffer containing 10 mM trimethylamino-ethanesulphonic acid (TES), *p*H 7 6 at 0 °C, 50 mM NaCl, 1 mM EDTA, 20 mM sodium molybdate and 10% glycerol (v/v) After the broken cell suspension had been centrifuged at 11,000g, aliquots of the cytosolic fraction were incubated with ~50 nM ^{3}H-triamcinolone acetonide (specific activity = 43 8 Ci per mmole) in the presence or absence of 2 μM triamcinolone acetonide to determine glucocorticoid receptor binding capacity[18]

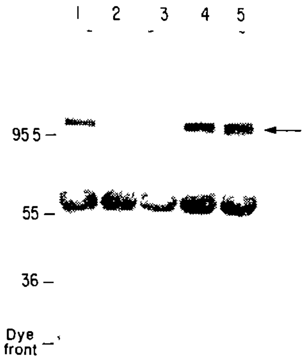


Fig 2 Cellular localization of the glucocorticoid null receptor Samples shown are cytosolic (lane 1) and nuclear (lane 2) fractions of WEHI-7 cells removed at time = 0 min, immediately after addition of DNP, cytosolic (lane 3) and nuclear (lane 4) fractions of WEHI-7 cells exposed to 1 mM DNP for 90 min, and Triton X-100-washed nuclear fraction of WEHI-7 cells exposed to 1 mM DNP for 90 min (lane 5) At various times after the addition of DNP, 1 ml of the WEHI-7 cell suspension was centrifuged at 400*g* for 1 5 min The cell pellet was lysed by a freeze-thaw procedure[10] in 1 ml of a lysis buffer containing 10 mM TES, *p*H 7 6 at 0 °C, 50 mM NaCl, 2 mM $MgCl_2$, 20 mM sodium molybdate and 10% (v/v) glycerol After lysis, the broken cell suspension was centrifuged at 400*g* at 3 °C for 3 min The supernatant was used without modification as the cytosolic fraction The nuclei were suspended in 5 ml of the lysis buffer, centrifuged at 400*g* for 5 min, and the supernatant discarded The nuclei were then resuspended in 600 μl of lysis buffer Nuclei to be treated with Triton X-100 were suspended in 5 ml of 0 05% Triton X-100 in 6 mM $MgCl_2$ for 5 min at 3 °C, centrifuged at 400*g* for 5 min, the supernatant discarded and the pellets resuspended in 600 μl of lysis buffer Glucocorticoid receptors were denatured in both fractions by addition of 200 μl of a detergent solution containing 10 mM TES, *p*H 7 6 at 0 °C, 10% (w/v) SDS, 10 mM sodium molybdate and 10% (v/v) glycerol, and heating the samples in a boiling water bath for 5 min (cytosolic fraction) or 10 min (nuclear fraction) After cooling to room temperature, 2 ml of a 5% solution of Triton X-100 (w/v) in 10 mM TES, *p*H 7 6 at 0 °C, 10 mM sodium molybdate, 10% (v/v) glycerol was added to bring the total volume of each sample to 5 ml Ascites fluid (10 μl) of BuGR2 antibody (directed against the glucocorticoid receptor[19]) was added to each sample, and the receptors purified on protein A-Sepharose CL-4B columns[10] Antibody-purified receptors were run on 7 5% polyacrylamide gels under denaturing and reducing conditions, transferred to nitrocellulose and identified by reaction with BuGR2 antibody and subsequent labelling with ^{125}I-sheep anti-mouse IgG (ref 10) After autoradiography the amount of receptor in each sample was determined by counting the ^{125}I in the receptor-containing region (arrow) The sizes (daltons × 10^3) and migration positions of prestained relative molecular mass standards (Diversified Biotech) are shown The immunoreactive protein which migrates just above the 55,000 standard is BuGR2 antibody

nucleus as the activated hormone/receptor complex and transducing hormonal activity, is released in the null form and the hormone binding form is subsequently regenerated It has been shown that the receptors for glucocorticoid and other steroids are phosphorylated and lose hormone-binding activity under cell-free conditions when treated with phosphatase (see refs 12 and 13 for reviews) Because the null receptor can also be generated in cells in the absence of steroid[3,6], it can directly interconvert with the normal unoccupied receptor Our demonstration that the null receptor is located in the nuclei of ATP-depleted cells, rather than in the cytosolic fraction, however, raises some interesting questions Among these is the possibility that the null receptor may be present as the predominant form in some of the mutant cells which have been considered receptor-deficient based on the lack of binding capacity or receptor protein in their cytosolic fractions[14,15]

We do not know where the null receptor is bound in the nucleus, nor whether it can initiate biological activity Many other questions also remain, such as how the null receptor is formed from the normal receptor, whether it interconverts with the normal receptor in metabolically intact cells, and whether its accumulation in nuclei of ATP-depleted cells means that release of receptors from nuclei is an energy-requiring step. Our identification and localization of the null receptor should allow definitive experimental answers to these and other questions.

We thank Dr Robert Harrison for the generous gift of the BuGR2 antibody used in these studies, and Sue Bagas for technical assistance. This research was supported by NIH grants AM 03535 and CA 17323, and by the Norris Cotton Cancer Center Core grant CA23108 from the US Public Health Service.

Received 2 September; accepted 8 October 1986.

1. Munck, A. & Brinck-Johnsen, T. *Excerpta Med. Int. Congr. Ser.* **132**, 472-481 (1967).
2. Munck, A. & Brinck-Johnsen, T. *J. biol. Chem.* **243**, 5556-5565 (1968).
3. Bell, P. A. & Munck, A. *Biochem. J.* **136**, 97-107 (1973).
4. Munck, A. *et al. J. Steroid Biochem.* **3**, 567-578 (1972).
5. Chader, G. J. *J. Neurochem.* **21**, 1525-1532 (1973).
6. Wheeler, R. H. *et al. J. biol. Chem.* **256**, 434-441 (1981).
7. Govindan, M. V. *Expl Cell Res.* **127**, 293-297 (1980).
8. Antakly, T. & Eisen, H. J. *Endocrinology* **115**, 1984-1989 (1984).
9. Welshons, W. V., Krummel, B. M. & Gorski, J. *Endocrinology* **117**, 2140-2147 (1985).
10. Mendel, D. B., Bodwell, J. E., Gametchu, B., Harrison, R. W. & Munck, A. *J. biol. Chem.* **261**, 3758-3763 (1986).
11. Sadowski, P. D. & Steiner, J. *J. Cell Biol.* **37**, 147-161 (1968).
12. Housley, P. R., Grippo, J. F., Dahmer, M. K. & Pratt, W. B. *Biochem. Actions Horm.* **11**, 347-376 (1984).
13. Rossini, G. P. *J. theor. Biol.* **108**, 39-53 (1984).
14. Bourgeois, S. & Gasson, J. C. *Biochem. Actions Horm.* **12**, 311-351 (1985).
15. Northrop, J. P., Gametchu, B., Harrison, R. W. & Ringold, G. M. *J. biol. Chem.* **260**, 6398-6403 (1985).
16. Stanely, P. E. & Williams, S. G. *Analyt. Biochem.* **29**, 381-392 (1969).
17. Khym, J. X. *Clin. Chem.* **21**, 1245-1252 (1975).
18. Munck, A. & Wira, C. *Meth. Enzym.* **36**, 255-264 (1975).
19. Gametchu, B. & Harrison, R. W. *Endocrinology* **114**, 274-279 (1985).

A new subunit of the human T-cell antigen receptor complex

Allan M. Weissman, Lawrence E. Samelson & Richard D. Klausner

Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

The T-cell antigen receptor binds antigen in association with a cell surface molecule encoded by the major histocompatibility complex (MHC)[1]. MHC restricted recognition of antigen by this receptor leads to the complex pattern of programmed gene expression that characterizes T-cell activation. The eventual understanding of human T-cell function will require the complete elucidation of the structure of the human T-cell antigen receptor. On human T cells, clonally determined, disulphide-linked α and β chains of the receptor are non-covalently and stoichiometrically associated with three additional polypeptides known as the T3 complex[2,3]. These receptor subunits are glycoproteins of relative molecular mass (M_r) 25,000 (25K) and 20K (γ and δ) and a non-glycosylated 20K protein (ε). Our studies of murine T cells show that the mouse T-cell antigen receptor consists of at least seven distinct polypeptide chains[4-6]. In addition to clonotypic α and β chains, the murine complex consists of glycoproteins of 26K and 21K and endoglycosaminidase F (endo F)-insensitive polypeptides of 25K, 21K and 16K. The latter, which we have termed ζ (zeta), exists as a homodimer within the complex. The 26K component (gp26) has been shown to be the murine analogue of the human δ chain[6,7]. Other cross species homologies remain to be established, however none of the described human receptor components appear similar to the murine ζ polypeptide. We report here the use of an antiserum raised against the murine ζ subunit to identify a previously unrecognized component of the human T-cell antigen receptor. This human protein is T-cell specific and biochemically similar to the murine ζ polypeptide.

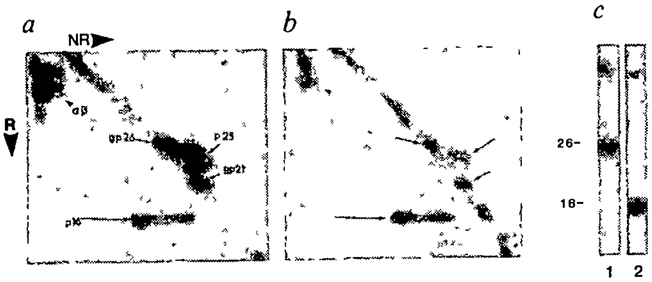


Fig. 1 Immunoprecipitation of radio-iodinated murine T-cell antigen receptor complex from the antigen specific hybridoma 2B4. Specific immunoprecipitates with *a*, the anti-clonotypic monoclonal antibody A2B4-2, or *b*, anti-ζ antiserum were subjected to two-dimensional diagonal non-reducing (NR)/reducing (R) electrophoresis. Arrows delineate subunits of interest. α–β are 40-45K glycoproteins, gp26 is a 26K glycoprotein (δ), p25 migrates at 22K under non-reducing and 25K under reducing conditions and gp21 is a 21K glycoprotein, p16 is the ζ chain. The 21K endo F insensitive protein, p21 (ref. 5) is not seen in this study. *c*, Solubilized cell membranes radio-iodinated using Iodobeads (Pierce) were immunoprecipitated and incubated with ionic detergents SDS and DOC before elution, reduction and SDS-PAGE. Lane 1, anti-δ rabbit serum, lane 2, anti-ζ rabbit serum.

Methods. 2B4 cells were maintained in exponential phase of growth as previously described[12]. Lactoperoxidase catalysed radio-iodination with ^{125}I at 4 °C, lysis of cells in 0.5% Triton X-100 and immunoprecipitation using antibodies precoupled to protein A Sepharose (Pharmacia) were performed as described[4]. 2B4 cell membranes were prepared and solubilized as described[5]. Incubation of immunoprecipitates in ionic detergents SDS and DOC was carried out by the addition of 0.2-0.4% SDS and 0.2% DOC to protein A Sepharose immunoprecipitates after an initial wash in 0.1% Triton X-100 wash buffer. After 15 min at 22 °C the immunoprecipitates were again washed in 0.1% Triton X-100 containing wash buffer and analysed after reduction on 12.5-13% polyacrylamide gels using the buffer system of Laemmli[16]. Two-dimensional diagonal electrophoresis was carried out using 12.5% SDS-PAGE tube gels in the first dimension on non-reduced material. After reduction and equilibration in 0.5% dithiothreitol, 0.1% SDS and 0.125 M Tris, *p*H 6.8 for 2 h at room temperature, SDS-PAGE in the second dimension was carried out on 12.5% polyacrylamide slab gels (Hoeffer) according to the manufacturer's recommendations. Fixing of gels, glycerol treatment, drying and autoradiography were all performed as previously described[4]. Solubilised 2B4 cell membranes were used for the preparation of antisera. Insoluble material was removed at 15,000g and the supernatant was immunoprecipitated with A2B4-2 covalently linked to cyanogen bromide-activated Sepharose (Pharmacia). After extensive washing of immunoprecipitates using buffers described previously[7], the receptor complex was eluted from the beads at 37 °C in a buffer consisting of 0.1% SDS, 50 mM Tris, *p*H 6.8, 0.5 M NaCl, 1 mM Phenylmethylsulphonylfluoride (Sigma), 10-100 µg ml^{-1} leupeptin (Boehringer Mannheim), 10 µg ml^{-1} Aprotinin (Boehringer Mannheim), 10 mM EDTA and 10 mM iodoacetamide. After elution the material was concentrated at 22 °C using Centricon 10 Microconcentrators (Amicon). Radio-iodinated, solubilized A2B4-2 membrane immunoprecipitates were added to the concentrated material, boiled in 3% 2-mercaptoethanol/10% glycerol and electrophoresed on 17.5% polyacrylamide gels cross-linked with Acrylaide (FMC Corporation). After autoradiography of the undried gel, the 16K band was cut from the gel, eluted in 0.1% SDS 10 mM Tris, *p*H 7.6 at 95 °C and concentrated as above. Material from approximately 1.5×10^{10} cell equivalents was emulsified in complete Freund's adjuvant (CFA) in a final volume of 2 ml and injected into an adult female New Zealand White rabbit by multiple subcutaneous injections. The animal was boosted after 5 weeks with antigen emulsified in CFA and bled two weeks thereafter.

Rabbit antisera were raised against the ζ chain after separation of receptor components by sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE). The antiserum was evaluated by immunoprecipitation of lysates from the murine T-cell hybrid 2B4. The proteins specifically immunoprecipitated from surface-iodinated cells by the anti-clonotypic monoclonal antibody A2B4-2 were compared to those precipitated by the

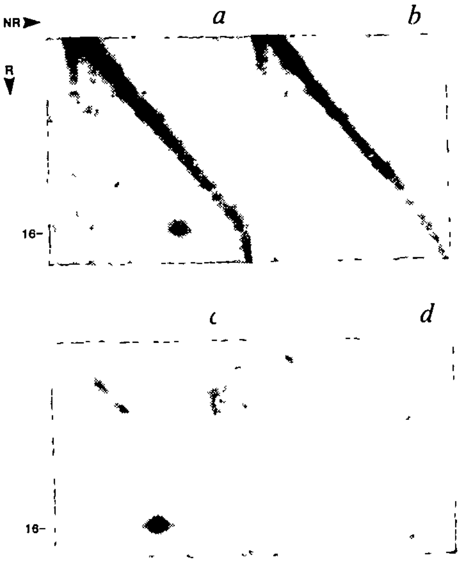


Fig 2 Identification of the human ζ chain Radio-iodinated solubilized membranes from Jurkat cells (*a*, *b*), human tonsillar T cells (*c*), and human tonsillar B cells (*d*) were immunoprecipitated with the anti-ζ serum (*a*, *c*, *d*) or non-immune rabbit serum (*b*) and subjected to electrophoresis on diagonal gels
Methods Jurkat human tumour cells were maintained as described previously[4] Purified tonsillar T and B cells (provided by Dr John Kehrl) were prepared as previously described[17] Tonsillar T cells were judged to be 90% pure by indirect immunofluorescence with the T-cell specific monoclonal antibodies OKT4 and OKT8 B cells were 94% surface immunoglobulin positive In this study 6 5 × 10^7 T cells and 5 × 10^7 B cells were used

anti-ζ serum on two-dimensional non-reducing/reducing diagonal gels (diagonal gels) (Fig 1*a*,*b*) The serum specifically precipitated all of the components of the T-cell receptor complex seen with the monoclonal antibody The relative intensities of the receptor components are different between panels *a* and *b* This may be due to disruption of the complex by the anti-ζ serum or to the presence of some dimeric ζ not associated with the receptor complex A 16K band present on the diagonal in the material precipitated by the antiserum may represent monomeric ζ A small amount of streaking is seen at 16K and 14K in both panels, this finding is variable and probably due to proteolysis This antiserum also immunoprecipitates an identical pattern of subunits from murine thymocytes (data not shown) In contrast, the anti-ζ serum fails to immunoprecipitate specific proteins from lysates of a murine B-cell tumour line LK 35 2, or from murine fibroblasts

To identify the direct target antigen of this antiserum we subjected radio-iodinated immunoprecipitates to an incubation in a buffer containing the ionic detergents SDS and deoxycholate (DOC) This treatment results in dissociation of the receptor complex[7] Immunoprecipitation from radio-iodinated membranes with an antiserum raised against the carboxy terminus of the murine δ chain[7] revealed a single band at 26K when the SDS/DOC incubation was carried out (Fig 1*c*) When the anti-ζ serum was used, SDS/PAGE under reducing conditions revealed a single prominent band at 16K In this experiment a small amount of material is seen at 50K with both of the specific antisera and with control serum

We used this anti-ζ serum to look for a similar protein in human T cells The human T-cell tumour line Jurkat expresses both α-β and the T3 complex on its surface[8] Solubilized membranes from these cells were radio-iodinated using lactoperoxidase and immunoprecipitates were electrophoresed on diagonal gels Immunoprecipitation with the anti-ζ serum (Fig 2*a*) revealed a 32K protein with a M_r of 16K after reduction No other components of the T-cell antigen receptor complex were detected with varying exposures of the autoradiograph Non-immune serum (Fig 2*b*) revealed a pattern of non-specific material on the diagonal similar to that seen in Fig 2*a*, but the 16K dimer was clearly not present Surface iodination of intact Jurkat cells gave identical results (data not shown) We prepared radioiodinated solubilized membranes from human tonsillar T and B cells to examine whether this protein is found in normal human T cells The anti-ζ-antiserum immunoprecipitates a 16K protein with a non-reduced M_r of 32K from T cells but not from B cells (Fig 2*c*,*d*) It is clear that in these cells no other components of the T-cell antigen receptor are specifically

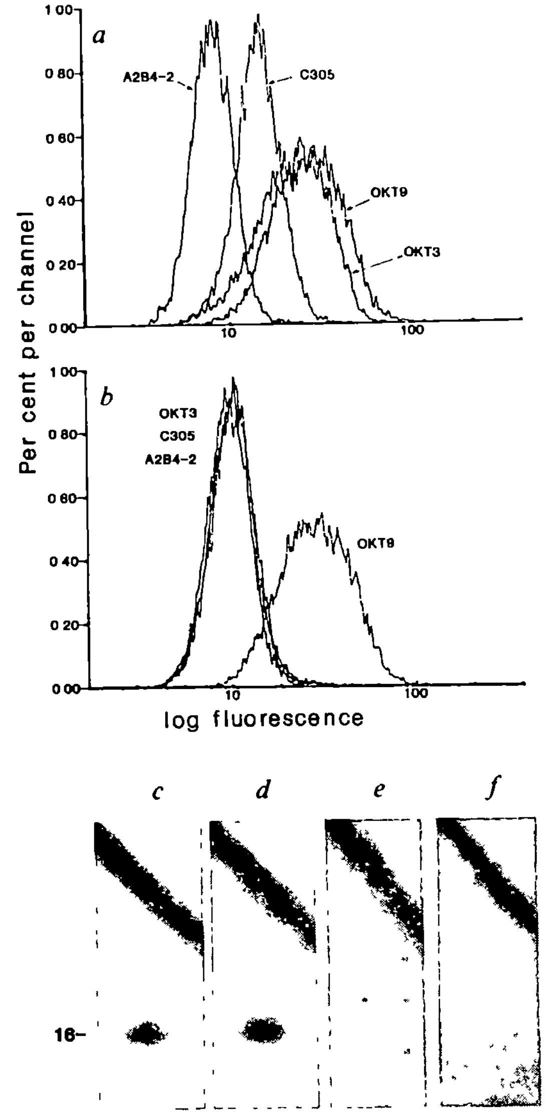


Fig 3 Jurkat cells incubated for 24 h with *a*, A2B4-2, or *b*, OKT3 ascites at a dilution of 1 1000 The cells were washed 3 times and treated for 30 min at 4 °C with ascites or culture supernatants of monoclonal antibodies A2B4-2, OKT3, OKT9 or C305 The cells were washed and incubated with fluorescein-conjugated Fab fragments from affinity-purified goat anti-mouse serum, then subjected to FACS analysis as described[18] Incubation with OKT9 for 24 h resulted in persistence of surface antibody as measured by FACS, interfering with analysis of antibody binding In parallel, Jurkat cells were surface iodinated after 24 h treatment with A2B4-2 (*c*), OKT9 (*d*) or OKT3 (*e*,*f*) Equivalent numbers of trichloroacetic acid precipitable counts were immunoprecipitated with anti-ζ serum (*c*, *d*, *e*) or non-immune serum (*f*) and subjected to electrophoresis on diagonal gels The position of ζ is shown The non-immune immunoprecipitate shown (*f*) is representative of similar control immunoprecipitations carried out after 24 h treatment with A2B4-2 or OKT9

precipitated with this antiserum Additional studies revealed that the 16 K human homodimer has the same apparent *p*I as the murine ζ chain and is likewise insensitive to endo F (data not shown)

Thus ζ, which is clearly a component of the murine T-cell antigen receptor complex, has a human analogue which (1) is immunologically related to the murine chain, (2) has identical size and charge characteristics to the murine protein and (3) is found in T cells but not B cells Based on these findings we analysed radio-iodinated immunoprecipitates of Jurkat cells obtained by using the anti-T3 monoclonal antibody OKT3 (ref 9) and the anticlonotypic antibody C305 (ref 8) With both antibodies, material that migrated identically to ζ was consistently seen on long exposures of autoradiographs of diagonal gels (data not shown) Although the level of ^{125}I associated with this material was significantly less than that found in the other subunits of the complexes, these results provided suggestive evidence that the human ζ polypeptide is part of the T-cell antigen receptor

One possible explanation for the apparently lower amount of ζ-like material in human receptor immunoprecipitates is that solubilization may lead to disruption of the complex We wished to determine whether the human ζ analogue was associated with the other receptor components on intact cells To test this, we used antibody-mediated co-modulation This approach has allowed the determination of specific non-covalent association of membrane proteins In particular it has been used to show the association of the T3δ chain with the clonotypic α-β heterodimer on the surface of human T cells[10] Modulation of surface T3 with anti-T3 reagents results in loss of both the clonotypic and T3 antigenic determinants without affecting other cell surface antigens Since the anti-ζ serum does not bind to the surface of cells as measured by fluorescence activated cell sorter (FACS) analysis, we modified the standard co-modulation technique by carrying out surface iodination of cells at 4 °C after antibody incubation This labelling allowed us to quantify surface ζ by immunoprecipitation

Jurkat cells were used for these studies because specific reagents for this cell line are available Jurkat cells were treated for 24 hours with ascites from hybridomas producing the OKT3, OKT9 or A2B4-2 monoclonal antibodies to modulate antigen receptors OKT9 is a monoclonal antibody which binds the human transferrin receptor[11] A2B4-2 does not bind to human cells[12] and served as a control When T3 surface expression was measured by FACS, the A2B4-2 pretreated cells bound considerably more OKT3 than did the OKT3 pretreated cells (compare Fig 3*a* with 3*b*) As expected, the α-β chains of the antigen receptor were co-modulated with the T3 antigen after OKT3 treatment as shown by the reduced binding of the anti-clonotypic C305 antibody Immunoprecipitation of the surface iodinated cells with C305 confirms the FACS data, essentially no residual heterodimer is seen after T3 modulation, but OKT9 and A2B4-2 treated cells have approximately the same amount of heterodimer (data not shown) The association of ζ with the T-cell antigen receptor complex on the intact cell is shown by the loss of ζ chain expression from the cell surface on OKT3-treated cells (Fig 3*c*-*f*) Thus ζ co-modulates with the α-β heterodimer and T3δ confirming the *in situ* association of these polypeptides The specificity of this co-modulation was demonstrated by FACS analysis OKT3-treated cells showed no loss of surface transferrin receptor binding compared to A2B4-2 treated cells, nor was there loss of binding of an anti-LFA-1 antibody MHM-23 (ref 13) (data not shown) In addition, when radio-iodinated OKT3-treated cells were compared to OKT9 or A2B4-2 treated cells on SDS-PAGE, few if any differences in the pattern of total cell surface proteins were seen

Solubilization of multicomponent membrane complexes invariably results in a certain degree of denaturation of native protein-protein interactions, as shown in studies of the T-cell antigen receptor and the IgE receptor[4,14] Optimizing ratios of lipid to detergent and the use of non-ionic detergents can minimize this denaturation[15] In the T-cell antigen receptor system however, ζ, which can be co-immunoprecipitated in a consistent stoichiometric relationship with other subunits from mouse cells, is not similarly co-immunoprecipitated from human T cells This disparity is probably due to minor structural differences that are manifest as varying degrees of disruption of subunit interactions upon solubilization

Aside from the identification of a new human T-cell antigen receptor component, we have shown that (1) when studying membrane protein complexes, comparison of the same receptor system in different species may allow for the identification of interactions that might not otherwise be detected, (2) because of the receptor denaturation that occurs on solubilization, methods other than co-immunoprecipitation, such as co-modulation and chemical cross-linking, may prove useful to help determine specific associations between cell surface structures

We thank Dr B J Fowlkes and Miss L Edison for assistance with fluorometry studies, Dr A Weiss for his gift of the antibody C305, Dr John Kehrl for providing tonsillar cells, Drs J Harford, Y Minami and J O'Shea for their help and advice, Miss H Luong for technical assistance, and Mrs K Perry for secretarial expertise in preparation of this manuscript

Received 17 June, accepted 25 September 1986

1 Marrack, P & Kappler, J *Adv Immun.* **38,** 1-24 (1986)
2 Borst, J , Prendiville, M A & Terhorst C *Eur J Immun.* **13,** 576-580 (1983)
3 Kanellopoulos, J M , Wigglesworth, N M , Owen, M J & Crumpton, M J *EMBO J* **2,** 1807-1814 (1983)
4 Samelson L E Harford, J B & Klausner, R D *Cell* **43,** 223-231 (1985)
5 Samelson, L E , Patel, M D , Weissman, A M , Harford, J B & Klausner, R. D *Cell* **46,** 1083-1090 (1986)
6 Oettgen, H C , Pettey, C L , Meloy, W L & Terhorst, C *Nature* **320,** 272-275 (1986)
7 Samelson L E Weissman, A M Robey, F A Berkower, I & Klausner, R D *J Immun* **137,** 3254-3258 (1986)
8 Weiss, A & Stobo, J D *J exp Med.* **160,** 1284-1299 (1984)
9 Kung, P C , Goldstein, G , Reinherz, E L & Schlossman, S F *Science* **206,** 347-349 (1979)
10 Meuer, S C *et al. J exp Med* **157,** 705-719
11 Schneider, C , Sutherland, R , Neuman, R & Greaves, M *J biol. Chem* **257,** 8516-8522 (1982)
12 Samelson, L E , Germain, R N & Schwartz, R H *Proc. natn. Acad. Sci. U S A.* **80,** 6972-6976 (1983)
13 Hildreth, J , Gotch, F , Hildreth, P & McMichael, A *Eur J Immun* **13,** 202-208 (1983)
14 Rivnay, B & Metzger, H *J biol. Chem.* **257,** 12800-12808 (1982)
15 Rivnay, B , Wank, S A , Poy, G & Metzger, H *Biochemistry* **21,** 6922-6927 (1982)
16 Laemmli, U K *Nature* **227,** 680-685 (1970)
17 Falkoff, R. J M , Zhu, L P & Fauci, A S *J Immun* **129,** 97-102 (1982)
18 Samelson, L E & Schwartz, R H *J Immun* **131,** 2645-2650 (1983)

Athymic mice express a high level of functional γ-chain but greatly reduced levels of α- and β-chain T-cell receptor messages

Yasunobu Yoshikai, Marciano D. Reis & Tak W. Mak

The Ontario Cancer Institute, and Departments of Medical
Biophysics and Immunology, University of Toronto,
500 Sherbourne Street, Toronto, Ontario, Canada M4X 1K9

T lymphocytes differentiate and mature in the thymus. It is here that thymocytes with reactivities to self antigens are eliminated and those with specificities to 'altered' self-major histocompatibility complex (MHC) gene products are positively selected[1]. The selections are presumably carried out on the basis of their T-cell antigen receptors (TcR). The genes of the α- and β-chain T-cell antigen receptors have been cloned[2–7]. A third T-cell specific gene capable of undergoing somatic rearrangement has also been identified[8]; the role of this third gene is not known. An order of expression of γ, β, then α is found during T-cell ontogeny[9,10]. But although α- and β-chain messages are often functional[11–15],

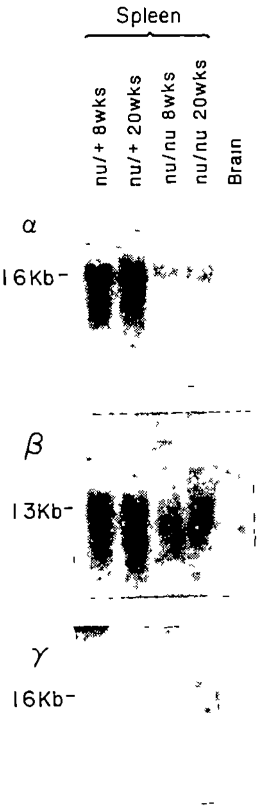


Fig. 1 Northern blot analysis of spleens in young and old nude mice Congenitally athymic nude (*nu*/*nu*) mice on a BALB/c (10th backcross) background and their littermate (*nu*/+) controls were used under pathogen-protected conditions in the animal facility maintained by the Ontario Cancer Institute, Toronto Original breeding pairs were obtained from Taconic Farms, N Y The life span of nude mice was 1 to 2 years under the maintenance conditions used All nude mice used in these studies appeared in good health RNA was extracted from spleens by the guanidine isothiocyanate and $CsCl_2$ gradient centrifugation procedures[29] and treated with glyoxal Total RNA (30 μg per spleen) was electrophoresed through 1% agarose in 10 mM sodium phosphate buffer *p*H 7 0 and transferred to Gene Screen Plus (NEN) Hybridization was to ^{32}P-labelled nick-translated α (5/10-20 Dα, *Hin*dIII-*Eco*RI fragment), β (5/10-20Dβ, *Taq*I-*Eco*RI fragment) and γ (8/10-2γ1 1, *Ava*I-*Eco*RI fragment) constant probe[11] Following hybridization in 5×SSC, 5×Denhardt's 10% dextran sulphate and 0 1% SDS at 65 °C, blots were washed three times in 2×SSC, 0 1% SDS at 65 °C and once in 0 2×SSC, 0 1% SDS at 65 °C

γ transcripts are rarely functional in thymocytes or mature T cells[16–18]. To define the sequential order of expression of these genes further and to continue the search for a possible role for the TcR γ gene products, we investigated the expression of 'functional' α-, β- and γ-chain transcripts in young athymic mice. These mice express an undetectable amount (less than one in 8×10⁵ spleen messages) of 'full-length' α- and β-chain T-cell receptor transcripts, but an increased level of expression of 'full-length' γ chain messages Nucleotide sequence analysis of four γ complementary DNAs show that all four γ transcripts sequenced are functional These findings suggest that γ gene products may be important in a prethymic or extrathymic pathway and may represent a second type of T-cell recognition, possibly in a lineage in which α and β genes are not used.

To investigate the expression of the α, β and γ T-cell receptor messages in athymic mice, we extracted RNA from total spleen cells from mice with genotype BALB/c (*nu*/*nu*) at 8 weeks and 20 weeks of age Control spleen RNA was also obtained from mice with genotype (*nu*/+) at 8 and 20 weeks of age The level of expression of these T-cell receptor genes was evaluated using Northern blots A typical result is shown in Fig 1 The experiments clearly show that a much lower level of the α- and β-chain T-cell receptor transcripts are found in both 8 and 20 weeks old athymic mice However the same level of expression of γ-chain messenger RNA is found in BALB/c (*nu*/*nu*) mice and in control mouse spleen Although the expression of these messages did not increase in old (20 weeks) control *nu*/+ mice, a higher level of these messages is found in old athymic mice The sizes of these γ-chain transcripts also suggest that they mainly consist of 'full length' 1 6 kilobase (kb) mRNA We have constructed cDNA libraries derived from spleen cells of 8 weeks old *nu*/+ mice and 8 weeks old athymic *nu*/*nu* mice The cDNA clones with sequences homologous to the constant region of either α-, β- or γ-chain genes were isolated and counted Results of one such experiment (Table 1) confirm that very low levels of α- and β-chain messages are found in the athymic mice whereas the level of γ chain messages is near normal

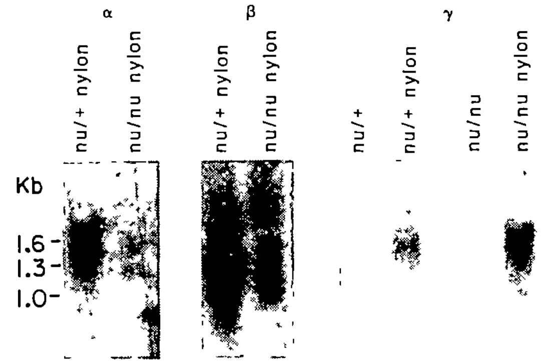


Fig 2 Northern blot analysis of nylon wool-passed spleen cells from nude mice Spleen cells were passed over nylon wool columns according to the method of Julius *et al*[20] Recovery of spleen cells after nylon wool passage varied (15–20%) Total RNA (15 μg each) and blots were prepared as described in Fig 1 Hybridization was carried out in 5×SSC, 5×Denhardt's, 10% dextran sulphate, 0 1% SDS at 65 °C Nick-translated constant probes of α, β and γ chain were used[33] After hybridization, blots were washed three times in 2×SSC, 0 1% SDS at 65 °C and once in 0 2×SSC, 0 1% SDS at 65 °C

The above data reflect the total amount of α-, β- or γ-chain T-cell receptor gene transcripts in total spleen RNAs However, it is known that spleens from athymic mice have a low number of T cells (reflected by the level of Thy-1⁺ cells[19]) We enriched for Thy-1⁺ spleen cells from 8-week-old *nu*/*nu* spleens using nylon wool[20] and compared them with the nylon-wool purified cells from spleens of mice with genotype *nu*/+ at the same age (Fig 2) The cells from athymic mice express a much lower level of both the α- and β-chain TcR transcripts Note also that, although control T cells express a high level of full-length transcripts of the α and β messages (1 6 and 1 3 kilobases (kb), respectively), cells from the athymic mice do not contain a high proportion of full length message This is particularly evident in the β- chain transcripts, where we find a prominent band of truncated message at 1 0 kb Examination of the sizes of eight inserts from the cDNA libraries described also confirms that all the α and β chain transcripts are less than the full length of 1.3 kb for α and 1 0 kb for β (Y Y, unpublished observations) These truncated transcripts are presumably composed of only *D-J-C* or *J-C* sequences[21-23] Thus the decrease in α- and β-chain transcripts in *nu*/*nu* mice is not solely due to the lower numbers of Thy-1⁺ cells

A different and interesting picture is seen for γ-chain messages Figure 2 clearly shows that after nylon-wool enrichment more γ-chain message can be found T cells from athymic mice express a several times higher level of γ-chain messages The majority of the γ cDNAs are 1 6 kb and are presumably 'full length'

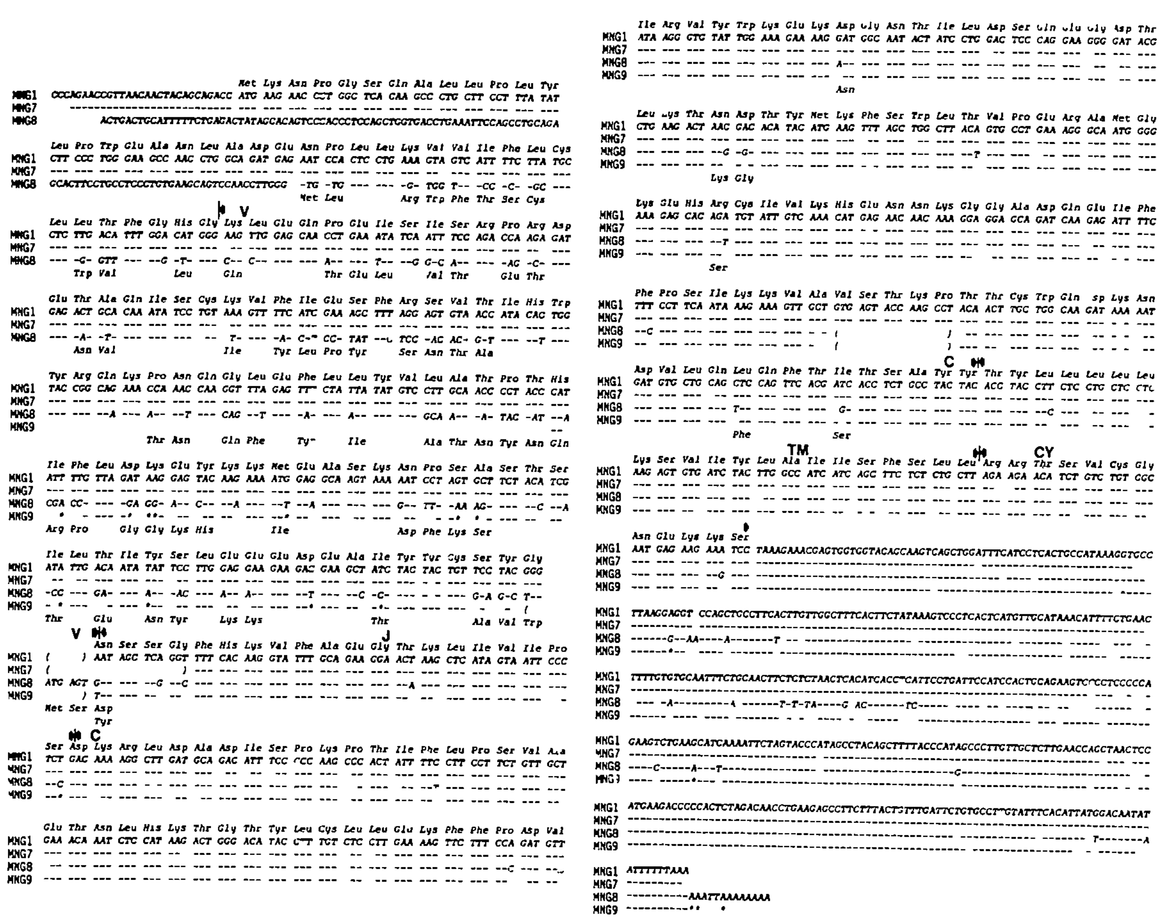


Fig. 3 The nucleotide and deduced amino acid sequences of four γ-gene cDNA clones isolated from *nu*/*nu* spleen. The cDNA inserts were subcloned into the M13mp9 bacteriophage, and their sequences were determined using both the universal M13 primer and the specific-primer-directed dideoxynucleotide sequencing technique[31,32]. The predicted amino acid sequence of cDNA clone MNG1 is shown above the nucleotide sequence. MNG8 and MNG9 amino-acid residue positions that differ from MNG1 are shown below the nucleotide sequence. The V, J, C transmembrane (TM) and cytoplasmic (CY) regions are indicated by the horizontal arrows. Dash, nucleotides in MNG7, MNG8 or MNG9 that are identical with those in MNG1; *, nucleotides in MNG9 that are identical with those in MNG8 ( ), missing nucleotides ( ).

To examine whether the γ-chain messages are functional, we sequenced four of these clones. The nucleotide sequences and deduced protein sequences are presented in Fig. 3. As can be seen, each of the four clones MNG1, MNG7, MNG8 and MNG9 contained sequences with open reading frames from the beginning of the variable region to the expected termination codons at the end of the constant regions. They also possess the entire *V*, *J* and *C* sequences presumably necessary for a functional receptor protein. The clones MNG1 and MNG7 use the *V*, *J* and *C* sequences of $V\gamma_4$ (3F9), $J\gamma_1$ (13.4), and $C\gamma_1$(13.4), while those of MNG8 and MNG9 correspond to $V\gamma_2$ (10.8A), $J\gamma_2$(10.5), and $C\gamma_2$(10.5) as defined in refs 24 and 34.

The above data clearly illustrate the difference between the expression of the α- and β-chain messages and the γ-chain transcripts in young athymic mice. The finding that there are full length, functional γ chain messages but no full length α and β transcripts in these athymic mice suggests that a complete thymic environment is not necessary for the rearrangement and expression of TcR γ gene.

Thy-1$^+$ T cells from athymic mice (purified using nylon wool) also contain a reduced level of both the α and β gene transcripts. This may be due to the presence of a different subpopulations of Thy-1$^+$ T cells in athymic spleens, possibly those of less mature T cells when compared to the normal mice. The finding that the α and β transcripts that do exist in young athymic mice are truncated is consistent with such a hypothesis. Another possibility is that the lower level of α and β TcR transcripts is due to a different environment in the athymic mice spleen or the predominance of a different subset(s) of splenic T cells in these *nu*/*nu* mice.

In either case it would be of interest to determine whether any *nu*/*nu* T cells use the α-β TcR heterodimer and whether a full wild-type germline repertoire of TcR α and β chain *V* and *J* gene segments are used in these athymic mice. This may require generation of 'functional' T cell clones from athymic mice. Some results have suggested that these cells may exist in *nu*/*nu* mice[25]. It would be of interest to examine whether older athymic mice use these α-β TcR genes.

It is interesting that although the levels of α- and β-chain TcR transcripts in athymic mice are dramatically reduced, the γ level in these mutant mice is elevated. The reason for this is not yet known. It is possible that the rearrangement and expression of γ-chain message occur before exposure to the thymus or is independent of the thymic environment, so that

Table 1 Expression of α, β and γ chain T cell receptor message in nude mouse spleen

	Frequency (%)	
	nu/+	*nu*/*nu*
α	0.006	0.0002
β	0.01	0.0004
γ	0.0006	0.0007

Double-stranded (ds) cDNA was synthesized from total RNA derived from spleens of *nu*/+ (8 weeks) and *nu*/*nu* (8 weeks) according to a modified procedure of Gubler and Hoffman[30]. After addition of the *Eco*RI linker, the cDNAs were cloned into λgt10. About 10^6 independent recombinant phages were plated on *E. coli* (C600 hfl) without amplification and screened using nick-translated constant probes of α, β and γ chain cDNA.

processing of these genes can occur outside the thymus Although the study of the ontogeny of the expression and rearrangement of the T-cell genes suggests an order of γ, β then α[9,10], these studies were done exclusively in the context of thymic environment Moreover, rearrangement of γ-chain genes before β-chain genes outside the thymus or in fetal thymus has not been documented It is possible that the rearrangement and expression of 'functional' γ-chain messages is more efficient than those of the α and β chain genes in the extrathymic pathways of T-cell differentiation in athymic mice

Perhaps the most exciting observation is that athymic mice contain a population(s) of T cells in which the γ-chain T-cell receptors are used It has been suggested that γ-chain gene products may be involved in the recognition of class I MHC gene products in a 'two receptor' model (one for antigen and one for MHC gene products)[9], but several recent findings suggest that the situation may be more complicated First, although it was thought that γ-chain genes are limited to cytotoxic T cells[26], helper T cells (especially autoreactive clones) have also been found to contain these messages[27] Secondly, Dembic *et al*[28] have reported the successful reconstitution of a new T cell specificity recognizing the antigen fluorescein in the context of murine class I D^d products by DNA transfer of only the α- and β-chain genes from a T cell with the same specificity This result suggests that the α-β heterodimer is sufficient for the recognition of both antigen and MHC gene products Finally, with only one exception, over thirty γ transcripts isolated from cytotoxic T lymphocytes[16,17,34], mature T lymphocytes and thymocytes[18] contain deletions, frameshift mutations, or unspliced messages Thus, it is likely that γ gene products may be important in a subpopulation or lineage of T cells responsible for a second type of immunorecognition that has yet to be identified Our observation that all four cDNAs obtained from athymic spleen cells are functional is in sharp contrast to the existing data on γ messages from functional T cells and thymocytes[16-18,34] The finding of no detectable full length α- and β-chain transcripts in young athymic mice suggests that the majority of these cells may not have the α-β heterodimer on their surfaces Athymic spleen cells should thus be the ideal cells for the study of the possible role of γ gene product They may contain populations of either pre-T cells or T cells in an extrathymic pathway, which may be involved in the establishment of tolerance or the recognition of self MHC gene products Other possibilities include γ gene products being receptors active in primitive defense mechanisms or receptors that recognize antigens and MHC gene products

This work was supported by grants from the MRC, NCI, NESERC of Canada and a special research grant from the University of Toronto Y Y is recipient of a fellowship from the MRC of Canada and M D R is a Terry Fox Fellow of the NCI of Canada We thank Drs Barry Toyonaga, Aikichi Iwamoto and Pam Ohashi for fruitful comments

Note added in proof A decrease of α-chain and β-chain messages in athymic mice[35] and the detection of γ-chain protein on the surface of human T cells have been reported[36-38]

Received 17 July, accepted 29 September 1986

1 Zinkernagel, R M & Doherty, P C *Adv Immun* **27,** 51-77 (1979)
2 Yanagi, Y *et al. Nature* **308,** 145-149 (1984)
3 Hedrick, S , Cohen, D , Nielson, E & Davis, M *Nature* **308,** 149-153 (1984)
4 Saito, H *et al. Nature* **312,** 36-40 (1984)
5 Chien, Y -H *et al Nature* **312,** 31-35 (1984)
6 Sim, G K *et al Nature* **312,** 771-774 (1984)
7 Yanagi, Y *et al Proc. natn Acad Sci U S A* **82,** 3430-3434 (1985)
8 Saito, H *et al. Nature* **309,** 757-762 (1984)
9 Raulet, D , Garman, R , Saito, H & Tonegawa, S *Nature* **314,** 103-107 (1985)
10 Snodgrass, H P , Dembic, Z , Steinmetz, M & von Boehmer, H *Nature* **315,** 232 233 (1985)
11 Barth, R K *et al Nature* **316,** 517-523 (1985)
12 Behlke, M A *et al Science* **229,** 566-570 (1985)
13 Arden, B , Klotz, J L , Siu, G & Hood, L E *Nature* **316,** 783-787 (1985)
14 Yoshikai, Y , Kimura, N , Toyonaga, B & Mak, T W *J exp Med* **164,** 90-103 (1986)
15 Kimura, N *et al. J exp Med* **164,** 739-750 (1986)
16 Rupp, F *et al. Nature* **321,** 876-878 (1986)
17 Reilly E B , Kranz D M , Tonegawa, S & Eisen H N *Nature* **321,** 878-880 (1986)
18 Yoshikai, Y *et al Fur J Immun* (in the press)
19 MacDonald, R H *et al J Immun* **126,** 865-870 (1981)
20 Julius, M H Simpson, F & Herzenberg, L A *Eur J Immun* **3,** 645-649 (1973)
21 Siu G *et al Nature* **311,** 344-348 (1984)
22 Clark, S P *et al Nature* **311,** 387-389 (1984)
23 Yoshikai Y *et al Nature* **312,** 521-524 (1984)
24 Heilig, J S & Tonegawa S *Nature* **322,** 836-840 (1986)
25 Hunig T in *T Lymphocytes Today* (ed Inglis, J R) 42-45 (Elsevier, Amsterdam, 1983)
26 Kranz, D M *et al Nature* **313,** 752-755 (1985)
27 Zuderer M Iwamoto, A & Mak, T W *J exp Med* **163,** 1314-1318 (1986)
28 Dembic, Z *et al Nature* **320,** 233-238 (1986)
29 Chirgmin J M Przybyla, A E , MacDonald, R J & Rutter, W J *Biochemistry* **18,** 5294-5299 (1979)
30 Gubler V & Hoffman B J *Gene* **25,** 263-269 (1983)
31 Sanger F , Nicklen, S & Coulson, A R *Proc natn Acad Sci U S A* **74,** 5463-5467 (1977)
32 Strauss, E C , Kobori J A , Siu, G & Hood, L *Analyt Biochem* **154,** 353-360 (1986)
33 Iwamoto, A *et al J exp Med* (submitted)
34 Iwamoto, A *et al J exp Med* **163,** 1203-1212 (1986)
35 Owen M J *et al Eur J Immun* **16,** 875-878 (1966)
36 Brener, M *et al Nature* **322,** 145-149 (1986)
37 Bank I *et al Nature* **322,** 179-181 (1986)
38 Weiss A , Newton, M & Crommie, D *Proc natn Acad Sci U S A* **83,** 6988-7002 (1986)

Homology between P-glycoprotein and a bacterial haemolysin transport protein suggests a model for multidrug resistance

James H. Gerlach, Jane A. Endicott, Peter F. Juranka,
Graham Henderson*, Farida Sarangi,
Kathryn L. Deuchars & Victor Ling†

The Ontario Cancer Institute, Princess Margaret Hospital and
The Department of Medical Biophysics, University of Toronto,
500 Sherbourne Street, Toronto, Ontario M4X 1K9, Canada

Increased expression of P-glycoprotein, a plasma membrane glycoprotein of relative molecular mass (M_r) 170,000 (170K), occurs in a wide variety of cell lines that exhibit pleiotropic resistance to unrelated drugs[1-4]. The presence of P-glycoprotein in human cancers refractory to chemotherapy suggests that tumour cells with multidrug resistance can arise during malignant progression[5]. We have discovered striking homology between P-glycoprotein and the HlyB protein, a 66K *Eschericia coli* membrane protein required for the export of haemolysin (protein of M_r 107K). P-glycoprotein can be viewed as a tandem duplication of the HlyB protein. The hydropathy profiles of the two proteins are similar and reveal an extensive transmembrane region resembling those found in pore-forming plasma membrane proteins The C-terminal region of P-glycoprotein and the HlyB protein contain sequences homologous to the nucleotide-binding domains of a group of closely related bacterial ATP-binding proteins We propose a model for multidrug resistance in which P-glycoprotein functions as an energy-dependent export pump to reduce intracellular levels of anticancer drugs

A search of the GenBank DNA Sequence Database (using the FASTN program from the National Biomedical Research Foundation (NBRF) Protein Identification Resource) detected extensive homology between the 3′ end of a P-glycoprotein complementary DNA clone (pL20, see Fig 1 legend) and the *hlyB* gene of *E coli* This gene is one of four involved in the production and transport of haemolysin *hlyA* encodes the α haemolysin molecule (107K), *hlyC* encodes a small (20K) cytoplasmic protein that may be involved in activating haemolysin, whereas *hlyB* and *hlyD* encode membrane proteins required for the export of α haemolysin from the bacterium[6] (for reviews see refs 7-10)

When the deduced amino-acid sequences of the *hlyB* gene[6] and pL20 are compared graphically, extensive C-terminal homology is evident (Fig 1*a*) Whereas there is little direct

---

* Present address Cangene Corp ,3403 American Drive, Mississauga, Ontario L4V 1T4, Canada
† To whom correspondence should be addressed

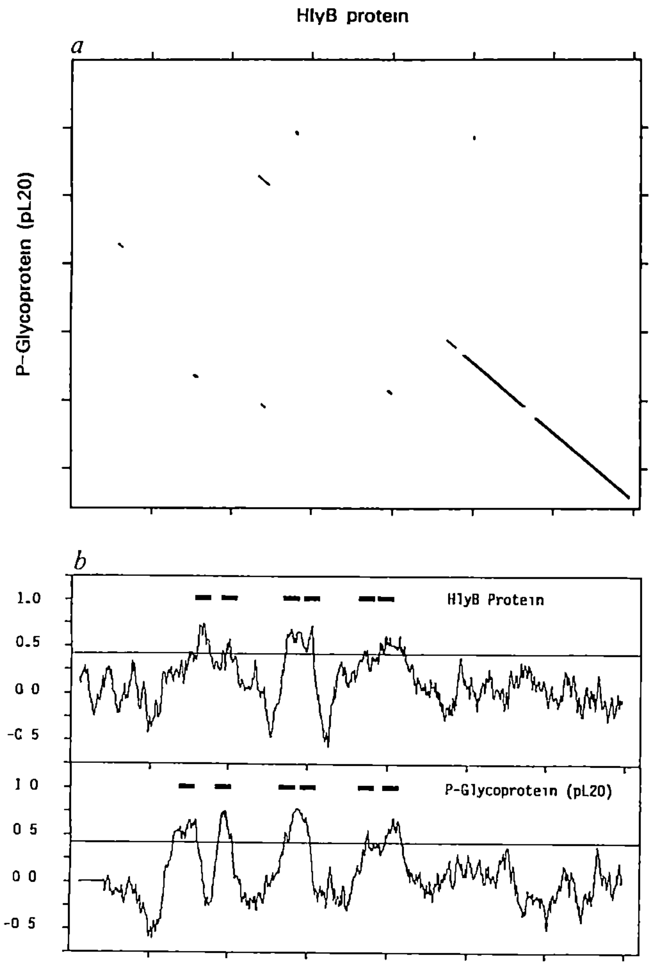


Fig. 1 Comparison of P-glycoprotein (C-terminal 655 amino acids) and the HlyB protein (707 amino acids). *a*, Segment comparison plot generated by the DOTMATRIX program from the National Biomedical Research Foundation (NBRF) Protein Identification Resource using the mutation probability data matrix for amino-acid replacements for scoring[35]. The window size was 21 amino acids and the minimum score required to generate a point was 30. Each division on the axes corresponds to 100 amino acids. *b*, Hydropathy profiles of the HlyB protein (top) and P-glycoprotein (bottom) generated using the normalized hydrophobicity values of Eisenberg *et al.*[11] and a window size of 21 amino acids. Regions above the line at 0.42 are candidates for membrane-spanning segments. The six predicted transmembrane segments in each sequence, as identified using the criteria of Eisenberg *et al.*[11], are indicated by solid bars. The mean hydrophobicity for 21 amino acid segments is indicated on the vertical axis. Divisions on the horizontal axis correspond to 100 amino acids.

Methods. The pCHP1 cDNA probe[31] was used to select a 2,321-base pair (bp) cDNA clone (pL20) from an Okayama and Berg pcD vector-library prepared from a drug-sensitive Chinese hamster ovary cell line as described[36]. For DNA sequencing, pL20 DNA was digested with *Pst*I or *Bam*HI and cloned into M13mp9. Both strands were sequenced using synthetic oligonucleotide primers and the dideoxy method[37]. The pL20 clone contains a 1,965-bp open reading frame followed by a 355-bp untranslated region and a poly(A) tail (sequence available on request). The 640-bp sequence from positions 1,167 to 1,806 is homologous to pCHP1.

homology in the remainder of the amino-acid sequence, hydropathy profiles of the two sequences are remarkably similar (Fig. 1*b*). Both proteins have a large hydrophobic region containing six potential transmembrane segments arranged in pairs. The hydrophobicities and the hydrophobic moments[11,12] of these putative transmembrane segments are consistent with channel-forming functions. The C-terminal region of P-glycoprotein has been localized to the cytoplasmic side of the plasma membrane with monoclonal antibodies[3]. Consequently, the hydropathy plots indicate that only the short stretches of amino acids joining paired transmembrane segments are exposed on the cell surface.

Sequence data from overlapping cDNA clones[13] indicate that the P-glycoprotein gene consists of a tandem duplication (unpublished observations), a structure confirmed in a functional, full-length mouse cDNA clone (P. Gros, personal communication). This, coupled with the protein structural analysis (Fig. 1), implies that P-glycoprotein is a near-perfect tandem duplication of the HlyB protein.

Only three breaks in the P-glycoprotein C-terminal domain are required to produce an excellent alignment with the 230 C-terminal amino acids of the HlyB protein (Fig. 2). Of the 228 possible amino acid matches, 107 (46.9%) are identical and a further 63 (27.6%) are conserved substitutions. This suggests these two proteins share both a common function and a common history. A search of the NBRF Protein Sequence Database (using the FASTP program) identified two other bacterial transport proteins, the HisP protein of the histidine permease system of *Salmonella typhimurium*[14] and the MalK protein of the maltose and maltodextrin transport system of *E. coli*[15], with significant homology with this region. These two proteins have homology with each other[16] as well as with the PstB protein of the phosphate-specific transport system of *E. coli*[17] and the OppD protein of the oligopeptide permease of *S. typhimurium*[18]. These four proteins are membrane-associated components of high-affinity, substrate-binding dependent, active transport systems (for a review see ref. 19). Homology between P-glycoprotein and these proteins is greatest in two regions (indicated A–A and B–B in Fig. 2) that are involved in nucleotide (probably ATP) binding and are thought to energize transport directly[18].

The extent of the protein sequence homology between P-glycoprotein, the HlyB protein, the HisP/MalK/PstB/OppD proteins and five other nucleotide-binding proteins is shown by cluster analysis (Fig. 3). The highest homology occurred between P-glycoprotein and the HlyB protein. Furthermore, the HisP/MalK/OppD cluster is more similar to the P-glycoprotein/HlyB cluster than to the bacterial OppD protein. This suggests both P-glycoprotein and the HlyB protein couple energy production to transport in a manner analogous to the binding-protein dependent bacterial transport systems. Although the cluster analysis does not give direct evolutionary relationships, the high degree of amino-acid sequence homology suggests that these proteins arose from a common ancestor and homology has been maintained by selection pressure resulting from functional similarities in these transport systems. Such a degree of evolutionary conservation between a eukaryotic protein (excluding mitochondrial and chloroplast proteins) and a bacterial protein is unprecedented, and raises the question of ancestry. Based on codon usage, it has been suggested[6] that *E. coli* has only recently acquired the chromosomal haemolysin gene group from a distantly related organism. It is therefore possible that *E. coli* has acquired at least part of the *hlyB* gene from a mammalian P-glycoprotein gene.

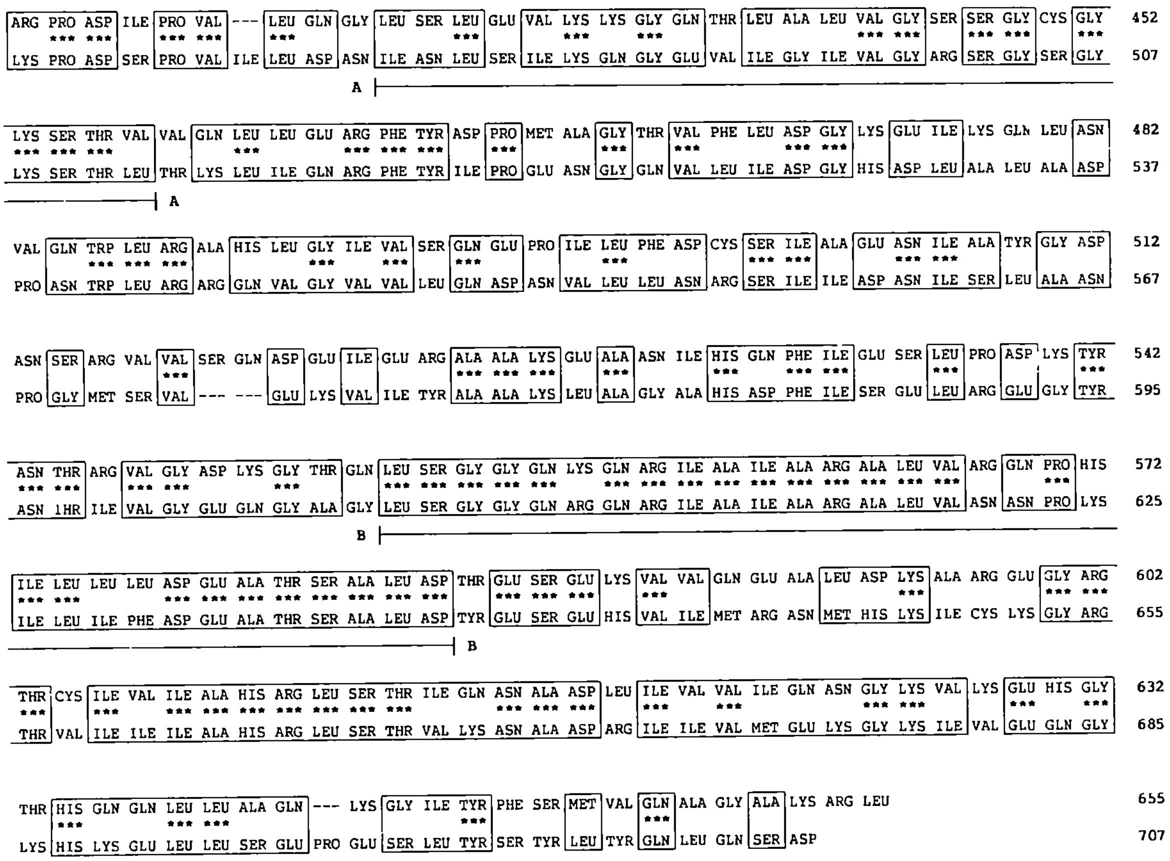

ARG PRO ASP ILE PRO VAL --- LEU GLN GLY LEU SER LEU GLU VAL LYS LYS GLY GLN THR LEU ALA LEU VAL GLY SER SER GLY CYS GLY 452
LYS PRO ASP SER PRO VAL ILE LEU ASP ASN ILE ASN LEU SER ILE LYS GLN GLY GLU VAL ILE GLY ILE VAL GLY ARG SER GLY SER GLY 507
A
LYS SER THR VAL VAL GLN LEU LEU GLU ARG PHE TYR ASP PRO MET ALA GLY THR VAL PHE LEU ASP GLY LYS GLU ILE LYS GLN LEU ASN 482
LYS SER THR LEU THR LYS LEU ILE GLN ARG PHE TYR ILE PRO GLU ASN GLY GLN VAL LEU ILE ASP GLY HIS ASP LEU ALA LEU ALA ASP 537
A
VAL GLN TRP LEU ARG ALA HIS LEU GLY ILE VAL SER GLN GLU PRO ILE LEU PHE ASP CYS SER ILE ALA GLU ASN ILE ALA TYR GLY ASP 512
PRO ASN TRP LEU ARG ARG GLN VAL GLY VAL VAL LEU GLN ASP ASN VAL LEU LEU ASN ARG SER ILE ILE ASP ASN ILE SER LEU ALA ASN 567
ASN SER ARG VAL VAL SER GLN ASP GLU ILE GLU ARG ALA ALA LYS GLU ALA ASN ILE HIS GLN PHE ILE GLU SER LEU PRO ASP LYS TYR 542
PRO GLY MET SER VAL --- --- GLU LYS VAL ILE TYR ALA ALA LYS LEU ALA GLY ALA HIS ASP PHE ILE SER GLU LEU ARG GLU GLY TYR 595
ASN THR ARG VAL GLY ASP LYS GLY THR GLN LEU SER GLY GLY GLN LYS GLN ARG ILE ALA ILE ALA ARG ALA LEU VAL ARG GLN PRO HIS 572
ASN THR ILE VAL GLY GLU GLN GLY ALA GLY LEU SER GLY GLY GLN ARG GLN ARG ILE ALA ILE ALA ARG ALA LEU VAL ASN ASN PRO LYS 625
B
ILE LEU LEU LEU ASP GLU ALA THR SER ALA LEU ASP THR GLU SER GLU LYS VAL VAL GLN GLU ALA LEU ASP LYS ALA ARG GLU GLY ARG 602
ILE LEU ILE PHE ASP GLU ALA THR SER ALA LEU ASP TYR GLU SER GLU HIS VAL ILE MET ARG ASN MET HIS LYS ILE CYS LYS GLY ARG 655
B
THR CYS ILE VAL ILE ALA HIS ARG LEU SER THR ILE GLN ASN ALA ASP LEU ILE VAL VAL ILE GLN ASN GLY LYS VAL LYS GLU HIS GLY 632
THR VAL ILE ILE ILE ALA HIS ARG LEU SER THR VAL LYS ASN ALA ASP ARG ILE ILE VAL MET GLU LYS GLY LYS ILE VAL GLU GLN GLY 685
THR HIS GLN GLN LEU LEU ALA GLN --- LYS GLY ILE TYR PHE SER MET VAL GLN ALA GLY ALA LYS ARG LEU 655
LYS HIS LYS GLU LEU LEU SER GLU PRO GLU SER LEU TYR SER TYR LEU TYR GLN LEU GLN SER ASP 707

Fig. 2 Aligned sequences of the C-terminal 232 amino acids of P-glycoprotein (upper sequence) and the C-terminal 230 amino acids of the HlyB protein (lower sequence) The alignment was generated by the ALIGN program from the NBRF Protein Identification Resource using the mutation probability data matrix for amino-acid replacements with a bias of two and a break penalty of six for scoring[35] Amino acids identical in both sequences are indicated by *** Conserved amino-acid substitutions (scores greater than zero in the mutation probability data matrix for amino-acid replacements[35]) are boxed The numbers at the ends of lines indicate the position in the sequence of the last amino acid in that line Three gaps introduced for alignment (single amino-acid breaks following positions 429 and 640 in the pL20 sequence and a two amino-acid break at position 572 in the HlyB sequence) are indicated (- - -) The underlined regions (A–A and B–B) contain the putative nucleotide-binding site and are highly conserved in other bacterial nucleotide-binding proteins (see Fig 3) Region A–A, characterized by the sequence hhG1/mSG(C/S)GKST ('h' indicates a hydrophobic amino acid conserved in P-glycoprotein, the HlyB protein and the HisP/MalK/PstB/OppD proteins), is glycine-rich and capable of forming a flexible loop that may undergo conformational change either on substrate binding or because of interaction with another protein domain[38] Region B–B contains four consecutive hydrophobic amino acids (which probably exclude water from the bound nucleotide to reduce hydrolysis[38]) followed by two acidic amino acids (hhhhDE) The high homology between P-glycoprotein and the bacterial transport proteins extends over 32 amino acids in this region suggesting an additional function(s) As this is the most conserved amino-acid sequence among the four binding-protein dependent transport proteins, it may couple energy production directly to transport

Based on homology and structural similarities between P-glycoprotein and the HlyB protein and on the homology with the four bacterial binding-protein dependent transport proteins, we propose that P-glycoprotein functions in the plasma membrane to export molecules from the cell It couples energy in the form of ATP directly to transport and effects transport through a channel formed by the transmembrane segments (Fig 4)

P-glycoprotein may export drugs directly (Fig 4*a*) This would require that the protein should not only be able to bind specifically many structurally diverse drugs but also that it should be able to release these drugs to the extracellular medium Support for such binding ability is provided by the vinblastine photoaffinity labelling of a 170K glycoprotein (P-glycoprotein?) in membrane vesicles from multidrug-resistant cells[20,21] The requirement for P-glycoprotein to recognize these compounds is eliminated if P-glycoprotein exports a protein rather than drugs (Fig 4*b*) Drugs bound by this protein (not necessarily reversibly) would be removed when the protein was exported Such binding might occur through hydrophobic domains such as those present in α haemolysin[6] We are examining culture medium from multidrug-resistant cells for such a protein

If P-glycoprotein functions in multidrug resistance in a manner analogous to that of the HlyB protein in haemolysin export, it is reasonable to expect these processes to have common aspects When inhibitors of energy production (for example potassium cyanide, 2,4-dinitrophenol, sodium azide) are added to mammalian cells in the absence of energy sources such as glucose, the net accumulation of drug is increased[22-26] This effect is greater in multidrug-resistant cells and, when glucose is restored, the intracellular concentrations of drug decrease more rapidly than in drug-sensitive cells An energy-dependent, efflux pump for drugs has been suggested as the mechanism of multidrug resistance[22,25,27] These compounds also inhibit the transport of haemolysin[28] Additionally, local anaesthetics (such as procaine), reduce the level of resistance of multidrug-resistant cells[29,30] and inhibit the export of haemolysin[28]

These findings have implications for cancer chemotherapy Although the homology reported here is for Chinese hamster P-glycoprotein, the conservation of P-glycoprotein, as judged by monoclonal antibodies[3] and DNA hybridization[31], makes it virtually certain that this homology will extend to other mammalian (including human) P-glycoprotein molecules It may be possible to screen for compounds to overcome multidrug

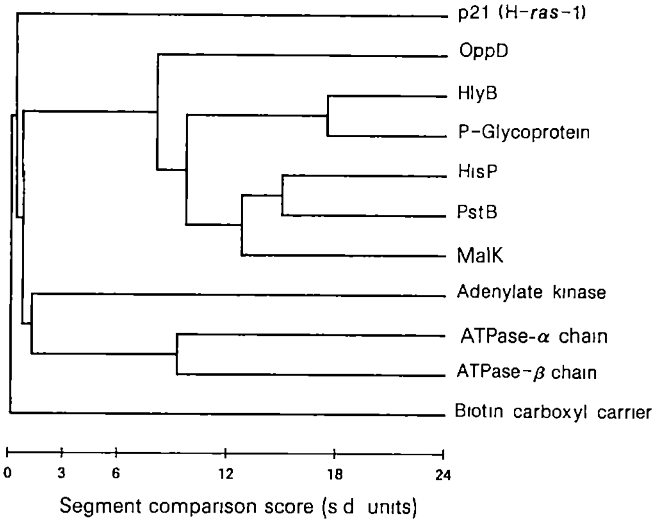


Fig. 3 Cluster analysis of the protein sequence homology of some nucleotide-binding proteins The RELATE program from the NBRF Protein Identification Resource[35] was used to determine objectively similarities between P-glycoprotein (pL20), the HlyB protein and nine proteins with nucleotide-binding regions homologous to regions A-A and B-B in Fig 2 Besides the HisP, MalK, PstB and OppD proteins, adenylate kinase (rabbit), the H-*ras*-1 p21 transforming protein (human), the biotin carboxyl carrier protein from *Propionibacterium shermanii* and the α and β chains of the *E coli* F_1-ATPase[38,39] were examined. These proteins were selected from known nucleotide-binding proteins because each has both of the nucleotide-binding regions present in the bacterial transport proteins and because they represent a functionally diverse group[38] The amino-acid sequences were from the NBRF Protein Sequence Database For each pair of proteins, the homology scores between all possible 25 amino-acid segments of one protein with all possible 25 amino acid segments in the other protein were determined using the mutation probability data matrix for amino-acid replacements[35] A numerical value derived from the distribution of scores was then determined for the actual sequences of the two proteins and for 500 runs performed using computer-generated randomized sequences of both proteins The difference between the value for the actual sequences and the mean value for the randomized runs was then divided by the standard deviation of the values from the randomized runs to yield a segment comparison score with units of standard deviation (s d) The probability of obtaining such a score by chance can then be estimated from a cumulative standardized normal distribution table This analysis has the advantage of being relatively insensitive to differences in the lengths of the proteins being compared[35] Clusters were formed by merging proteins with the highest segment comparison scores As clusters were formed, the score used for merging was the weighted average of the scores of the members of a cluster Note that cluster analysis forces the eventual merging of all proteins Only statistically significant groups should be considered representative of structural homology Thus we consider only the group consisting of P-glycoprotein, the HlyB protein and the HisP/MalK/PstB/OppD proteins and the group containing the ATPase α and β chains to be significant (the probability of obtaining a score >6 s d is $<10^{-9}$ whereas the probability of a score >10 s d is $<10^{-23}$)

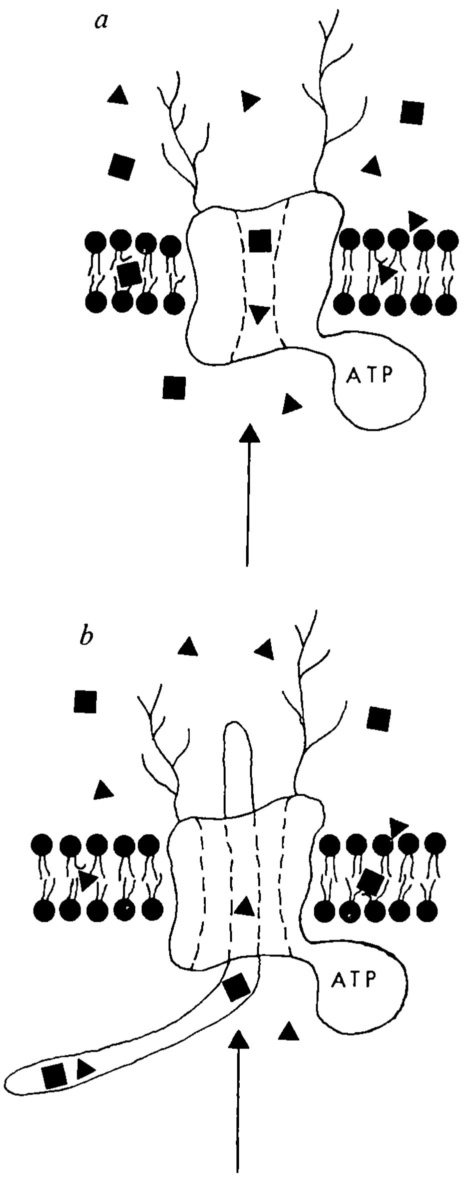


Fig 4 Models for the role of P-glycoprotein in multidrug resistance In *a* P-glycoprotein exports two structurally different drugs (■, ▲) directly through a channel In *b* the drugs are bound by an unidentified protein that is then exported from the cell by P-glycoprotein The presumptive ATP-binding site (ATP) is on the cytoplasmic side of the plasma membrane Drugs are shown entering the cell by diffusion through the lipid bilayer Arrows indicate the direction of movement through the membrane channel For clarity, only half the P-glycoprotein tandem duplication is shown In keeping with known channel-forming proteins[12], several P-glycoprotein molecules may be required to form this channel

resistance based on their ability to inhibit haemolysin export In a similar manner, it may be possible to screen existing and potential anticancer agents for their differential ability to kill haemolytic and non-haemolytic bacteria In this context, it is intriguing that there is a brief report that a nalidixic acid-resistant strain of *E coli* had also become haemolytic[32]

Although we have presented a model for the role of P-glycoprotein overexpression in multidrug resistance, the function of P-glycoprotein in drug-sensitive cells is not clear It may serve to protect cells from low levels of toxic lipophilic compounds and its role in multidrug resistance could simply be an extension of this function in response to increased concentrations of such compounds Alternatively, P-glycoprotein may have a transport function unrelated to multidrug resistance and only upon gross exaggeration of this function (for example by gene amplification) does resistance become apparent If this is the case, then the most likely candidates for transport by P-glycoprotein are, by analogy with haemolysin transport, large proteins One group of potentially haemolysin-like proteins is the pore-forming molecules released by cytotoxic cells such as eosinophils, cytotoxic T lymphocytes and NK-like cells[33,34] These molecules appear to be well conserved, in keeping with the conserved nature of P-glycoprotein

The potential existence of a previously unknown protein-translocation mechanism in eukaryotic cells is perhaps the most important aspect of the homology between P-glycoprotein and the HlyB protein The high degree of conservation of P-glycoprotein suggests this protein plays a fundamental role in mammalian cells As P-glycoprotein appears to be encoded by a multigene family[31], there may be a number of homologous systems, each translocating a different protein or proteins

Supported by the National Cancer Institute (NCI) of Canada and by NIH Grant CA37130 J A E is a recipient of a Medical Research Council (MRC) of Canada Studentship, P F J is an MRC Fellow and K L D is a recipient of an NCI Studentship We thank Dr G Bradley for helpful discussion and Drs D Trichler and S Minkin for advice on cluster analysis

Note added in proof The homology between the binding-protein dependent transport proteins and the HlyB protein has been independently reported[40] The nucleotide-binding region of these proteins also has homology with bacterial proteins involved in cell division[40], nodulation during nitrogen fixation[40] and DNA repair[41]

Received 22 August, accepted 15 October 1986

1 Juliano, R. L & Ling, V *Biochim biophys Acta* **455**, 152-162 (1976)
2 Kartner, N, Riordan, J R. & Ling, V *Science* **221**, 1285-1288 (1983)
3 Kartner, N, Evernden-Porelle, D, Bradley, G & Ling, V *Nature* **316**, 820-823 (1985)
4 Gerlach, J H, Kartner, N, Bell, D R & Ling, V *Cancer Surveys* **5**, 25-46 (1986)
5 Bell, D R., Gerlach, J H, Kartner, N, Buick, R N & Ling, V *J clin Oncol* **3**, 311-315 (1985)
6 Felmlee, T, Pellet, S & Welch, R. A *J Bact* **163**, 94-105 (1985)
7 Cavalieri, S J, Bohach, G A & Snyder, I S *Microbiol Rev* **48**, 326-343 (1984)
8 Hacker, J & Hughes, C *Curr Topics Microbiol Immun* **118**, 139-162 (1985)
9 Hartlein, M *et al J cell Biochem* **22**, 87-97 (1983)
10 Mackman, N, Nicaud, J M, Gray, L & Holland, I B *Current Topics in Microbiol Immun.* **125**, 159-181 (1986)
11 Eisenberg, D, Schwarz, E, Komaromy, M & Wall, R *J molec Biol* **179**, 125-142 (1984)
12 Eisenberg, D *A. Rev Biochem* **53**, 595-623 (1984)
13 Van der Bliek, A M, Van der Velde Koerts, T, Ling, V & Borst, P *Molec cell Biol* **6**, 1671-1678 (1986)
14 Higgins, C F *et al Nature* **298**, 723-727 (1982)
15 Gilson, E, Nikaido, H & Hofnung, M *Nucleic Acids Res* **10**, 7449-7458 (1982)
16 Gilson, E, Higgins, C F, Hofnung, M & Ames, G F L *J biol Chem* **257**, 9915-9918 (1982)
17 Surin, B P, Rosenberg, H & Cox, G B *J Bact* **161**, 189-198 (1985)
18 Higgins, C F, Hiles, I D, Whalley, K & Jamieson, D J *EMBO J* **4**, 1033-1040 (1985)
19 Ames, G F L *A Rev Biochem* **55**, 397-425 (1986)
20 Safa, A R, Glover, C J, Meyers, M B, Biedler, J L & Felsted, R L *J biol Chem* **261**, 6137-6140 (1986)
21 Cornwell, M M, Safa, A R, Felsted, R L, Gottesman, M M & Pastan, I *Proc natn Acad Sci U S A* **83**, 3847-3850 (1986)
22 Dano, K. *Biochim biophys Acta* **323**, 466-483 (1973)
23 See, Y P, Carlsen, S A, Till, J E & Ling, V *Biochim biophys Acta* **373**, 242-252 (1974)
24 Skovsgaard, T *Cancer Res* **40**, 1077-1083 (1980)
25 Inaba, M, Kobayashi, H, Sakurai, Y & Johnson, R. K *Cancer Res* **39**, 2200-2203 (1979)
26 Inaba, M, Fujikura, R & Sakurai, Y *Biochem Pharmacol* **30**, 1863-1865 (1981)
27 Lalande, M E, Ling, V & Miller, R G *Proc natn Acad Sci U S A* **78**, 363-367 (1981)
28 Springer, W & Goebel, W *J Bact* **144**, 53-59 (1980)
29 Carlsen, S A, Till, J E & Ling, V *Biochim biophys Acta* **455**, 900-912 (1976)
30 Chlebowski, R. T, Block, J B, Cundiff, D & Dietrich, M F *Cancer Treat Rep* **66**, 121-125 (1982)
31 Riordan, J R *et al Nature* **316**, 817-819 (1985)
32 Walton, J R. & Smith, D H *J Bact* **98**, 304-305 (1969)
33 Young, J D-E, Peterson, C G B, Venge, P & Cohn, Z A *Nature* **321**, 613-616 (1986)
34 Young, J D-E, Cohn, Z A & Podack, E R *Science* **233**, 184-190 (1986)
35 Dayhoff, M O *Atlas of Protein Sequence and Structure* Vol 5, Suppl 3 (National Biomedical Research Foundation, Washington, 1978)
36 Elliott, E M, Okayama, H, Sarangi, F & Henderson, G *Molec Cell Biol* **5**, 236-241 (1985)
37 Sanger, F, Nicklen, S & Coulson, A R *Proc natn Acad Sci U S A* **74**, 5463-5467 (1977)
38 Fry, D C, Kuby, S A & Mildvan, A S *Proc natn Acad Sci U S A* **83**, 907-911 (1986)
39 Walker, J E, Saraste, M, Runswick, M J & Gay, N J *EMBO J* **1**, 945-951 (1982)
40 Higgins, C F *et al Nature* **323**, 448-450 (1986)
41 Doolittle, R F *et al Nature* **323**, 451-453 (1986)

Mature murine B lymphocytes immortalized by Kirsten sarcoma virus

A. H. Lichtman*, D. S. Reynolds*,
D. V. Faller† & A. K. Abbas*

* Department of Pathology, Harvard Medical School and
Brigham and Women's Hospital, Boston, Massachusetts 02115, USA
† Dana-Farber Cancer Institute and Harvard Medical School,
Boston, Massachusetts 02115, USA

Clonal, antigen-specific, functionally responsive cell populations have proved critical for the analysis of the activation and regulation of lymphocytes. Such studies with B lymphocytes, the precursors of antibody-secreting cells, are hampered by the difficulty in generating phenotypically mature, antigen-reactive lines from defined cell populations. One method is to use acutely transforming retroviruses, which can transform B-lineage lymphocytes *in vitro*. However, Abelson murine leukaemia virus (A-MuLV) infection of murine bone marrow cells *in vitro* yields mostly immature B-cell lines[1–5], and infection of murine bone marrow cells with murine sarcoma viruses carrying *ras* related genes produces only immature lymphoid cell lines[6]. Retroviruses which contain *ras* can immortalize nonlymphoid cells without causing loss of mature phenotypic characteristics[7]. We used *ras*-containing Kirsten sarcoma virus (KiSV) pseudotyped with an amphotropic MuLV helper virus, to infect a purified population of mature, hapten-binding murine splenic B lymphocytes, aiming to generate mature B-cell lines to use as models for the study of B-cell growth and differentiation physiology. Immortalized B-cell lines which retain the same mature phenotype as the starting population, including hapten-specific binding, were produced. This is the first demonstration of a method for immortalizing selected antigen-binding B lymphocytes, and the first example of immortalization of mature B cells *in vitro* with an acutely transforming *ras*-containing retrovirus.

In most reported studies of retroviral transformation of B-lineage lymphocytes *in vitro*, the starting cell populations are heterogeneous, immature haematopoietic precursors With suitable manipulation of culture conditions, surface immunoglobulin positive lines are obtained[3] A-MuLV infection of clonal pre-B cells gives rise to B-cell lines which express surface immunoglobulin[4] These transformed lines are of unknown antigen specificity and are often autonomously growing and tumorigenic, or depend on nonphysiological stimuli such as high concentrations of lipopolysaccharide for continued growth[5] We infected mature B lymphocytes, selected on the basis of antigen-specific receptor expression, with an acutely transforming retrovirus To maximize the efficiency of infection of such phenotypically differentiated cells, we used viral pseudotypes composed of the transforming retrovirus and an amphotropic helper virus We isolated 2,4,6-trinitrophenol (TNP)-binding B cells (TNP-B cells) from the spleens of unimmunized (C3H × DBA/2)F$_1$ (C3D2F$_1$) mice by binding to and elution from TNP-gelatin plates This method normally enriches TNP-specific plaque-forming cells 50-100-fold and 50-90% of the cells form rosettes with TNP-modified sheep red blood cells[8] KiSV pseudotyped with amphotropic (virus 4070) or ecotropic (Moloney MuLV) helper virus was propagated in BALB/c 3T3 cells (BALB/3T3), as previously described[7] We used an amphotropic helper virus because acute transforming viruses tend to infect and immortalize only immature stem cells because of a limitation of target cell host range and this may be overcome using a helper virus with a broad range of cell/tissue infectivity[7,9,10] TNP-B cells were cocultivated with lethally irradiated BALB/3T3 cells producing amphotropic pseudotyped KiSV (KiSV-ampho) in microwell culture plates Growing colonies of lymphoblastoid cells were found in 6-25% of wells in 1-3 weeks (Table 1) Between 12-26% of wells gave rise to colonies in different experiments, consistent with each colony being clonal in origin by Poisson statistical analysis Assuming that the colonies of proliferating cells in each well were clonal, the efficiency of outgrowth of colonies was from 3×10^{-6} to $1\ 2 \times 10^{-5}$ Long-term cell lines were established from these colonies and have been growing continuously for 13 months The characteristics of two of these lines are described below No lymphoblastoid colonies were observed in cocultivations of TNP-B cells with BALB/3T3 cells producing ampho-MuLV

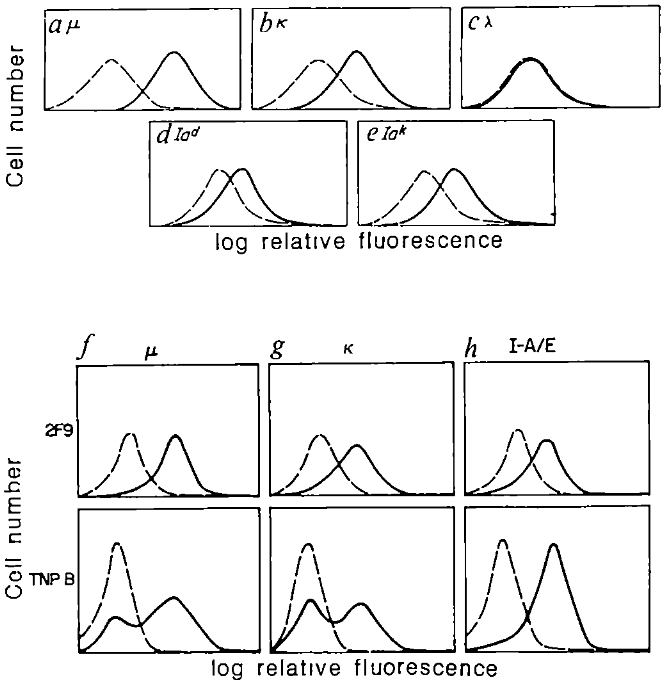


Fig 1 Surface immunoglobulin and Ia expression on a KiSV-immortalized B-cell line *a–e*, 2F9 cells were incubated with hybridoma supernatants specific for *a*, IgM (Bet 2) (ref 12), *b*, Ig κ light chain (187 1) (ref 13), *c*, Ig λ light chain (4/101) (from T Imanishi-Kari, MIT), *d*, I-A^d (MKD6) (ref 14), *e*, or I-A^k (10 2 16) (ref 15), for 30 min at 0 °C, followed by fluorescein-conjugated mouse anti-rat κ (FITC-MAR 18 5, Becton-Dickinson) (*a–c*), or fluorescein-conjugated goat anti-mouse IgG (*d*, *e*) for 30 min at 0 °C, *f–h*, 2F9 and freshly isolated C3D2F$_1$ TNP-binding B cells were incubated with Bet 2, 187 1 or M5/114 (rat anti-mouse I-A/E) hybridoma supernatants, followed by FICT MAR 18 5 Stained cells were fixed in 1% paraformaldehyde and examined by flow cytometry (solid traces) Dotted traces, nonspecific fluorescence (identical samples incubated with fluorescein-conjugated reagents alone) Note that 2F9 and TNP-B cells show comparable intensities of fluorescence, but the 2F9 cells are significantly larger (average diameter 12 5 μm compared to 6 4 μm for normal B cells) and, therefore, express IgM, κ and I-A/E at significantly lower densities Also, approximately 65% of the TNP-B cells are IgM, κ$^+$ (but 100% are Ia$^+$)—this may be due to immunoglobulin modulation during affinity purification or presence of Ia$^+$ non-B cells in this population

alone, or with uninfected BALB/3T3 cells TNP-B cells were also cocultivated with BALB/3T3 cells producing KiSV with Moloney-MuLV helper virus The titres of sarcoma virus produced by KiSV (ampho)- or KiSV (Moloney)-infected BALB/3T3 cells were similar, at approximately 5×10^7 focus forming units per ml In the cultures with KiSV (Moloney) 10% of wells showed colony growth, but no lines could be established from such colonies No viable cells could be recovered from cultures of TNP-B cells alone after 1 week

Two cell lines derived from KiSV infection of TNP-B cells (2F9 and 6H3) have been extensively characterized These lines have doubling times of 24 and 6–12 h, respectively, and grow as nonadherent cells in liquid culture medium with 10 or 20% fetal bovine serum and $2\ 5\times10^{-5}$ M 2-mercaptoethanol The phenotype of these cells was analysed by fluorescence-flow cytometry Both 2F9 and 6H3 express surface IgM with κ light chain (Fig 1*a*) In addition, they both express Ia determinants of the H-2^k and H-2^d haplotypes (Fig 1*a*) Compared to purified C3D2F$_1$ TNP-binding B cells, the KiSV-immortalized lines express lower densities of surface IgM and I-A/E determinants (Fig 1*b*)

The hapten specificity of 2F9 and 6H3 has been analysed by two methods First, up to 25% of these cells form rosettes with TNP-modified sheep erythrocytes (TNP-SRBC) Such rosette assays depend on the amount of hapten bound to the SRBC and the avidity of the surface receptors for the hapten (A K A , unpublished data) It is therefore not surprising that only a fraction of the cells formed rosettes No rosettes are seen in suspensions of these KiSV-immortalized B cells with unmodified sheep red blood cells, nor in suspensions of non-TNP-specific lines (such as the murine B-cell lymphoma A20) and TNP-SRBC (data not shown) The specificity of rosette formation was confirmed in inhibition assays using 2,4-dinitrophenyl (DNP)-lysine Inhibition of rosetting is 50% at a concentration of 6×10^{-6} M DNP-lysine with normal C3D2F$_1$ TNP-B cells and at $1\ 6\times10^{-5}$ M with the 2F9 line This suggests that 2F9 is an immortalized B cell whose receptors are of relatively low affinity, the population of normal hapten-binding B lymphocytes is heterogeneous Hapten binding was also analysed by flow cytometry Both 2F9 and 6H3 preferentially bind TNP-modified human serum albumin (TNP-HSA) over unmodified HSA, while a control BALB/c B lymphoma (A20 2J) that is not specific for TNP, shows no such preference (Fig 2) Thus the hapten specificity of the original uninfected B cells is retained in the immortalized lines

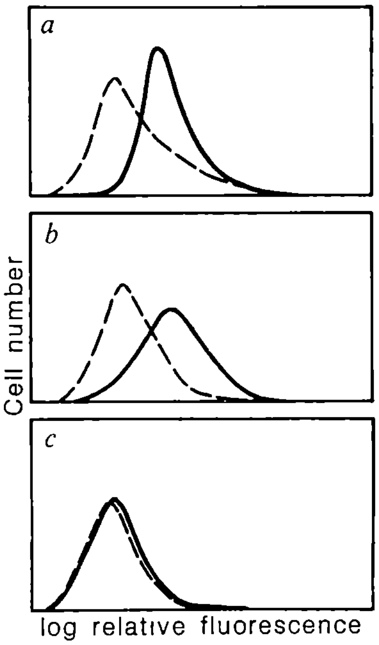


Fig 2 TNP-HSA binding to KiSV-immortalized B-cell lines Cells from the KiSV-immortalized B-cell lines 2F9 and 6H3 (*a*, *b*), and *c*, the murine B-cell lymphoma (ref 8) were incubated with dotted line, 20 μg ml^{-1} human serum albumin (HSA), solid line, TNP-conjugated HSA (TNP-HSA), for 30 min at 0 °C, followed by fluorescein-conjugated goat anti-HSA for 30 min, 0 °C Paraformaldehyde-fixed cells were then analysed by flow cytometry

Viral gene expression in the KiSV-immortalized B-cell lines was assayed by probing for viral transcripts Cytoplasmic RNA prepared from these cell lines was analysed for the presence of Ki-*ras* homologous RNA by dot blot hybridization using a v-*ras* probe derived from KiSV Significant levels were found in both 2F9 and 6H3 (Fig 3) Low or undetectable levels of v-*ras* RNA were present in other murine lymphoblastoid lines (for instance A20 2J) and in normal splenic TNP-B cells (Fig 3) Cytoplasmic RNA dot blot hybridizations using an amphotropic-MuLV *env* specific probe[7] also indicated a high level of MuLV-specific RNA in 2F9 and 6H3 (not shown) Thus, transcripts from both the helper virus and the sarcoma virus are present in the KiSV-immortalized lines Pelletable reverse transcriptase activity was measured in ultracentrifuged culture supernatants, by a standard technique in which [^{32}P]dTTP is incorporated into DNA using

Table 1 Generation of lymphoblastoid colonies by cocultivation of murine splenic B cells with virus-producing BALB/3T3 cells

Splenic B-cell target population	Virus (helper)	Positive wells	Estimated efficiency
TNP-binding (C3H×DBA/2)F_1	KiSV (ampho-MuLV)	43/250	0 86×10^{-5}
TNP-binding (C3H×DBA/2)F_1	KiSV (Moloney-MuLV)	5/50	0 5×10^{-5}
TNP-binding (C3H×DBA/2)F_1	Ampho-MuLV	0/50	0
TNP-binding (C3H×DBA/2)F_1	None	0/200	0

TNP-binding spleen cells were purified from (C3H×DBA/2)F_1 (C3D2F_1) mice by binding to and elution from TNP-gelatin plates These cells were mixed with γ-irradiated BALB/3T3 cells (10,000 rad) producing Kirsten sarcoma virus pseudotyped with either amphotropic or ecotropic (Moloney) murine leukaemia virus in RPMI 1640 supplemented with 20% fetal bovine serum, 2 0×10^{-5} M 2-mercaptoethanol and 5 μg ml^{-1} polybrene (Sigma) B cells (2×10^{-4}) were plated with 5×10^{-4} BALB/3T3 cells in 200 μl flat bottom microculture wells Cultures were fed with fresh medium every 2-4 days Wells were microscopically assayed for the presence of growing cells The number of microwells with growing lymphoblastoid clusters appearing within 1-6 weeks (positive wells) and the fraction of plated B cells induced to grow out into colonies, assuming a clonal origin of these colonies (estimated efficiency), are given Control cultures were plated with either no BALB/3T3 cells, or BALB/3T3 cells producing amphotropic MuLV alone The data are pooled from three separate experiments

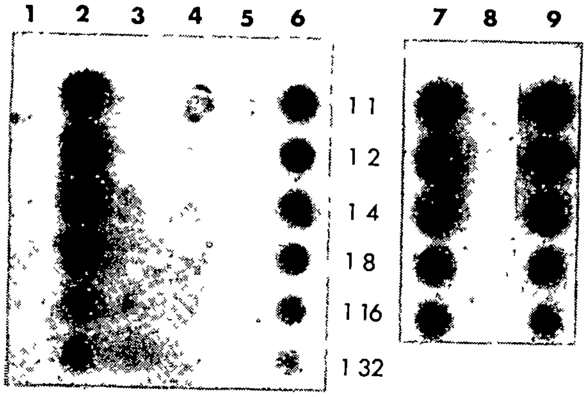


Fig. 3 RNA dot blot analysis of B-cell lines for Ki-*ras* specific RNA Ki-*ras* homologous cytoplasmic RNA was analysed in lanes 1 and 8, Abelson-MuLV-immortalized B-cell lines 1D12 and 12E11 (respectively), lanes 2 and 7, KiSV-immortalized B-cell lines 2F9 and 6H3 (respectively), lane 3, normal murine splenic B cells, lane 4, the murine B-cell lymphoma A20 2J, lane 5, Abelson-MuLV (amphotropic MuLV)-producing BALB/3T3 cells, lanes 6 and 9, KiSV (amphotropic-MuLV)-producing BALB/3T3 cells Cytoplasmic RNA was prepared from 4×10^6 cells of each type by NP-40 lysis and denatured in 5% formaldehyde at 55 °C for 15 min Samples were serially 2-fold diluted and blotted onto nitrocellulose with a vacuum manifold (Schleicher & Schuell) Blots were baked at 80 °C for 1 h under vacuum and hybridized to a ^{32}P-labelled Ki-*ras* specific probe (the 600 base-pair *SstII*/*HindII* fragment from KiSV) (ref 7) in 10% dextran, 5×SSC, at 42 °C, for 18 h The final washes of the blots were in 0 2×SSC at 50 °C Autoradiography was carried out at −80 °C for 18 h using intensifying screens[7,16]

a poly(A) template and trichloroacetic acid precipitable radioactivity is measured[17] Supernatant from cultures of 2F9 (10 ml) yielded 8,110 counts per minute (c p m) in this assay, compared to 677 c p m from an uninfected BALB/3T3 culture supernatant, and 128,761 c p m from the culture supernatant of a Mo-MuLV-producing BALB/3T3 line Thus 2F9 produces low levels of viral particles

We have shown here that introduction of the v-*ras* oncogene into mature B lymphocytes by retroviral infection can immortalize them This is the first demonstration of *ras*-induced immortalization of mature lymphocytes *in vitro* and it is likely that the use of an amphotropic helper virus was important to our success KiSV pseudotyped with an ecotropic helper (Moloney-MuLV) generated a few lymphoblastoid colonies none of which could be established as permanent lines (Table 1) The amphotropic envelope allows infection of cells that do not express membrane receptors for ecotropic viral envelope glycoproteins[9,10] The same approach has been used to immortalize differentiated human umbilical vein endothelial cells[7], and to infect mature helper T-cell lines *in vitro*[11] Mature B cells may have few or no functional receptors for ecotropic retroviruses, but do have amphotropic retroviral receptors

The cell lines described here retain their mature phenotype, and preliminary functional studies indicate that they can be further mitotically stimulated by reagents which stimulate normal B cells, particularly the helper T-cell-derived lymphokine, B-cell stimulatory factor 1 (data not shown) These cells do not secrete immunoglobulin constitutively, we are attempting to induce increases in immunoglobulin gene transcription, biosynthesis and/or secretion with activating stimuli, including lymphokines The technique of infecting purified, functionally well-defined lymphocyte populations with acute transforming retroviruses will be useful for the development of model systems to study lymphocyte physiology as well as the role of oncogene expression in lymphocyte growth deregulation

This work was supported by NIH grants AI22560 and CA37169 A L is supported by USPH Training Grant HL07627 D V F is a fellow of the John A Hartford Foundation D S R is supported by the PhD Program in Immunology, NIH grant CA09141

Received 16 April, accepted 18 September 1986

1 Risser, R *Biochim biophys Acta* **651**, 213-244 (1982)
2 Rosenberg, N & Baltimore, D *J exp Med* **143**, 1453-1463 (1976)
3 Rosenberg, N , Baltimore, D & Scher, C D *Proc natn Acad Sci U S A* **72**, 1932-1936 (1975)
4 Whitlock, C A , Ziegler, L J & Treiman, J L *Cell* **32**, 903-911 (1983)
5 Servnian, L A & Rosenberg, N *Molec Cell Biol* **6**, 183-194 (1986)
6 Pierce, J H & Aaronson, S A *J exp Med* **156**, 873-887 (1982)
7 Faller, D V *et al.*, in preparation
8 Rock, K. L , Benacerraf, B & Abbas, A K *J exp Med* **160**, 1102-1113 (1984)
9 Hartley, J W & Rowe, W P *J Virol* **19**, 19-25 (1976)
10 Rasheed, S , Gardner, M B & Chan, E *J Virol* **19**, 13-18 (1976)
11 Lichtman, A H , Abbas, A K. & Faller, D V , in preparation
12 Kung, J T , Sharrow, S O , Sieckmann, D G , Lieberman, R & Paul, W E *J Immun* **127**, 873-876 (1981)
13 Yelton, D E , Desaymard, C & Scharff, M D *Hybridoma* **1**, 5-11 (1981)
14 Kappler, J W , Skidmore, B , White, J & Marrack, P *J exp Med* **153**, 1198-1214 (1981)
15 Iverson, G M , Goldsby, R A & Herzenberg, L A *Curr Topics Microbiol Immun* **81**, 115-129 (1978)
16 Flyer, D C , Burakoff, S J & Faller, D V *J Immun* **135**, 2287-2292 (1985)
17 Baltimore, D E & Smoler, D *Proc natn. Acad Sci U S A* **68**, 1507-1511 (1971)

Erratum

Energy conversion catalysis using semiconducting transition metal cluster compounds

N. Alonso Vante & H. Tributsch

Nature **323**, 431-432

In this letter, the metal atoms in the Fig 1 legend were incorrectly defined They should read ●, Mo, ○, Ru, ⊜, Se

Stratigraphic record of earthquake incidence

CISNE[1] uses the Lofer cycles of the late Triassic Dachstein Formation (Northern Limestone Alps, Austria and Germany) as an illustration of how stratigraphy may record histories of stick-slip movement and thereby of earthquake incidence along fault zones As the discoverer of the emergence-submergence oscillations of the Lofer cycles, which I attributed to a sea-level oscillation[2], I am very interested in Cisne's alternative tectonic way of producing periodic emergence of a platform margin His model explains why the signs of emergence (microkarst, relict soils) should be restricted to the outer margin of the carbonate platform But it necessitates some revision of current sedimentological and/or tectonic concepts

It is necessary first of all to place these Triassic terranes into tectonic context They are now part of the giant Oberostalpin ('upper Austride') klippe that extends for 450 km, from the Swiss border to Vienna, in itself a patchwork of thrust sheets and klippen of various facies, complicated by high-angle faults, and developed in a complex and controversial history[3-7]

In Triassic time, the terrane was part of a tectonic trough, filled in the north with Triassic platform carbonates, while the southern part was open and accumulated the thin pelagic-hemipelagic Hallstatt beds[5]

In several places, the toe of the reef complexes interfingers with ammonite-bearing, pink or greenish calcilutites of Hallstatt type However, the bulk of the Hallstatt deep-water rocks are now found as discreet klippen masses emplaced on top of the platform facies and its deeper Jurassic/early Cretaceous cover

Cisne's stick-slip fault would have defined the southeastern margin of the platform, separating it from the Hallstatt facies It would have added to regional subsidence an additional component, either on the platform or on the Hallstatt side Throwing the Hallstatt side down, relative to the platform, accelerates the subsidence of the platform margin during the 'stick' phase, leading through intertidal into subtidal conditions The slip phase allows this margin to snap back into emergence, followed by truncation, and resubmergence in the next stick episode In the alternative case, throwing the platform down slows the platform margin's subsidence during the stick phase, and sedimentation leads to progressive shoaling, then the margin snaps down in the slip phase Thus, the Hallstatt-down model leads to upward deepening cycles, the platform-down model to upward shoaling cycles One of the most distinctive characteristics of the Lofer cycles is their generally upward-deepening character, which has seemed unusual to students of carbonate platforms If these cycles were caused by stick-slip faulting, then the motion was down on the Hallstatt side

It is here that difficulties develop If the Hallstatt basin underwent more subsidence than the platform, it must have subsided by more than the thickness of Norian platform sediments (some 1 5 km) during Norian time Since the Norian of the Hallstatt facies probably does not exceed a thickness of 150 m, the late Norian calls for a water depth of 1 5-2 km This depth is about the limit to which pelagic aragonite shells persist on the present sea floor (the aragonite compensation depth) It seems that modern compensation depths go back to the dramatic rise in plankton carbonate productivity in late Jurassic time[8] earlier carbonate compensation depths lay at far shoaler levels Thus, conventional wisdom assigns the Hallstatt carbonates with their ammonites to water depths of several hundred metres, not the required 2 km

I can think of three possible ways to escape this dilemma—all of which would require revision of current concepts One is that the oceans of the Triassic were saturated with calcium carbonate to much greater depths than those of the Jurassic, allowing preservation of aragonitic shells at depths of 1 5-2 km

A second possibility is that the Hallstatt basin was oceanic, a Triassic antecedent of the Jurassic (Pennine) ocean strip, and that most of the Hallstatt beds were seamount deposits, scraped off during southward subduction of that oceanic basin, and emplaced over the margin of the platform as it began to be subducted in mid-Cretaceous time

The third possibility is that the basin fronting the Loferite cycles was deep The Hallstatt beds were deposited in a shoaler part of the trough, to the west The Hallstatt facies was first brought to the longitude of the Lofer cycles by sinistral strike-slip displacements, probably in the mid-Jurassic when the Lammer-Torrenn zone underwent tens of kilometres of sinistral motion[5], and were then emplaced by Juvavian (mid-Cretaceous) thrusting

My preference is for the second of these models, which implies an 'accreted prism' origin for the entire stack of Alpine sediments

ALFRED G FISCHER

Department of Geological Sciences,
University of Southern California,
Los Angeles, California 90089-0741, USA

1 Cisne, J *Nature* **323**, 320-322 (1986)
2 Fischer, A G *Kansas geol Surv Bull* **169**, 107-149 (1966)
3 Schwarzacher, W *Jb geol Bundesanst, Wien* (1946)
4 Krystin L, Schaffer, G & Schlager W *Neues Jb Geol Palaont Abh* **137**, 284-304 (1971)
5 Fischer A G *Verh geol Bundesanst, Wien* **1/2**, 20-33 (1965)
6 Zankl, H *Geol Rdsch* **56**, 128-139 (1967)
7 Zankl, H in *Sedimentology of Parts of Central Europe* (ed Muller, G) 1471-85 (Kramer, Frankfurt, 1971)
8 Garrison, R E & Fischer, A G *Soc Econ Geol Paleontol spec Publ* **14**, 20-55 (1969)

Primary structure of the histidine-rich protein of *Plasmodium lophurae*

SIR—We wish to address the literature concerning the sequence of the histidine-rich protein (HRP) and the corresponding gene from the avian malaria parasite *Plasmodium lophurae* Following publication of a DNA sequence derived from a *P lophurae* genomic clone[1], we were forced to reassess our own data, which differed extensively from the 3′ half of the published coding sequence Our results were derived from the complementary DNA clones[2] that were used to isolate the genomic clone A corrigendum to the genomic sequence was published[3], in the revised sequence, 120 base pairs (bp) were deleted and 90 bp were inserted Our analyses[4] independently suggested that the published genomic sequence contained an insertion and a deletion, each of about 120 bp, together with a region of about 100 bp containing multiple insertions of one or two bp in the genomic sequence The region of the genomic sequence between the insertion and deletion sites was not revised[3] There is a significant difference in the amino acid compositions predicted by the two sequences, the most extreme being a fourfold difference in phenylalanine content

What then are the most probable structures of the HRP and its gene? Making the assumptions that the published 5′ DNA sequence[1] and corrections[3] to the numbers of downstream repeats not covered by the cDNA sequence are correct and that the 3′ sequence from the cDNA[4] should take preference, a composite sequence can be constructed The mature HRP (289 amino acids) predicted by this sequence has a molecular weight of 37,109 and contains 71 3% histidine, 7 2% proline, 7 0% alanine, 6 2% glutamic acid and smaller amounts of eight other amino acids This composition is consistent with three independent determinations of the composition of purified HRP (cited in ref 4)

DAVID O IRVING
GEORGE A M CROSS

The Rockefeller University,
New York,
New York 10021-6399, USA

1 Ravetch, J V, Feder, R, Pavlovec, A & Blobel, G *Nature* **312**, 616-620 (1984)
2 Wallach, M, Cully, D F, Haas, L O C, Trager, W & Cross, G A M *Molec biochem Parasit* **12**, 85-94 (1984)
3 Ravetch, J V, Feder, R, Pavlovec, A & Blobel, G *Nature* **317**, 558 (1985)
4 Irving, D O, Cross, G A M, Feder, R & Wallach, M *Molec biochem Parasit* **18**, 223-234 (1986)

Your Research Demands The Latest Resources

Williams & Wilkins means top-name international authority in medical and scientific periodicals
Get the best and the latest in clinical and experimental investigations from eminent researchers
Keep up with the cutting edge of your science

NEW

Your source for the best in leukemia research

LEUKEMIA

Official Journal of the Leukemia Society of America and the Leukaemia Research Fund, U K

Editors-in-Chief

Sv.-Aa. Killmann, M.D., Professor of Medicine, Rigshospitalet, University Hospital of Copenhagen, Blegdamsvej 9, DK-2100 Copenhagen, Denmark

C.N. Muller-Bérat, M.D., Department of Clinical Immunology, Statens Seruminstitut, Amager Boulevard 80, DK-2300 Copenhagen, Denmark

Deputy Editor, **Melvin F. Greaves,** United Kingdom,
Associate Executive Editor, **Julio E. Celis,** Denmark,
Honorary Editor, **Jean Bernard,** France

Associate Editors

D Van Bekkum, C.D Bloomfield, A Burny, C.M Croce, E P Cronkite, M Dexter, E J Frenich, R.C Gallo, K.D McCredie, D Metcalf, K Takatsuki, T Waldmann

Advisory Board

P J Fialkow, A Graessmann, W.A Haseltine, D R Lowy, J.D Rowley

Editorial Board

J Boniver, R Bravo, J Breton-Gorius, D Catovsky, Z Darzynkiewicz, S Dube, G E Francis, E Frindel, G Gaedicke, R.P Gale, R.E Gallagher, D W Golde, T Graf, J D Griffin, A Hagemeijer, F G J Hayhoe, N Haran-Ghera, J N Ihle, G Janossy, J Kersey, W Knapp, T W LeBien, A.M Mauer, R.P McCaffrey, E McCulloch, J Minowada, F Mitelman, W C Moloney, R.L Momparler, K Nilsson, W Ostertag, M Seligmann, P C Vincent, F.A Vyth-Dreese, B Westermark, P.H Wiernick, F Wong-Staal, R Zittoun

LEUKEMIA is your monthly resource for current, comprehensive, and open-to-debate coverage on all facets of leukemia, associated illnesses, and normal hemopoiesis From *clinical studies to molecular hematology and cell biology,* you'll read top-quality, peer-reviewed articles targeted to meet your interests in leukemia within

- molecular biology
- virology
- genetics
- pharmacology
- therapeutics
- bone marrow transplantation
- pathology
- and other allied fields

Today's top researchers, physicians, and educators publish their best work in **LEUKEMIA**. *accepted and scheduled manuscripts are published within three months, and rapid communications,* ***Leukemia Letters,*** *within eight weeks*

Get **LEUKEMIA**—a springboard for your own contributions to research in the field

Monthly/Vol 1 (January 1987)/ISSN 0887-6924
Individual $90/yr. ($105/yr outside U S)
Institution $340/yr ($360/yr outside U S)

Contributors Investigators wishing to submit articles should contact either Editors-in-Chief.

NEW

A journal devoted to the molecular mechanisms of cellular regulation

MOLECULAR ENDOCRINOLOGY

A Journal of The Endocrine Society

Editor-in-Chief

E. Brad Thompson, M.D.; University of Texas Medical Branch, Galveston, Texas, U S.A.

Associate Editors

Joel F Habener, Anthony R Means, Gordon M Ringold, Michael G Rosenfeld

Editorial Advisory Board

Stuart A Aaronson, F Carter Bancroft, Mariano Barbacid, Miguel Beato, Suzanne Bourgeois, Henry Bourne, Darrell H Carney, Pierre Chambon, P Michael Conn, Priscilla S Dannies, Daryl K Granner, Bernd Groner, Jan-Ake Gustafsson, Tony Hunter, Michael Karin, Marc E Lippman, Richard E Mains, Harold L Moses, Malcolm G Parker, Jeffrey M Rosen, William J Rutter, Herbert H Samuels, E Aubrey Thompson, Jr , Lewis T Williams

This new peer-reviewed, interdisciplinary journal from The Endocrine Society supplies a forum for the increasing number of scientists working on the molecular mechanisms of hormone action **MOLECULAR ENDOCRINOLOGY** publishes the best original articles on molecular mechanisms of cellular regulation, including studies involving the investigation of cloned genes and gene fragments relevant to hormone action, receptor and second messenger systems at the molecular level, and closely related areas such as lymphokines and oncogenes Other areas include

- autocrine controls
- genetic engineering
- structural organization of genes
- cell cycle
- regulation of gene expression
- excitable membrane channels
- control of cell motility
- chromatin & nuclear scaffold

Get monthly updates on every development in research laboratories around the world Take this opportunity to get **MOLECULAR ENDOCRINOLOGY.**

Monthly/Vol 1 (January 1987)/ISSN 0888-8809
Individual $90/yr. Institution $145/yr In-training $50/yr. (add $20/yr outside U S)

Contributors: Investigators wishing to propose or submit articles appropriate for **MOLECULAR ENDOCRINOLOGY** should contact the Editor-in-Chief at this address Basic Science Bldg , Room 226 F36, University of Texas Medical Branch, Galveston, Texas 77550 U S.A.

It's easy to order. Use the coupon below or contact your subscription agent or bookseller In the U S , call toll-free 1-800-638-6423 (except from Alaska) In Maryland, call collect 528-4105

clip & mail

I want the latest research resources.

clip & mail

Send **LEUKEMIA**
☐ Individual (U S $90, Outside U S $105)
☐ Institution (U S $340, Outside U S $360)

Send **MOLECULAR ENDOCRINOLOGY**
☐ Individual (U S. $90 Outside U S $110)
☐ Institution (U.S $145, Outside U S. $165)
☐ In Training (U S $50, Outside U S $70) **include institution/status*

Orders outside U.S must be prepaid, U.S dollars only Maryland residents add 5% sales tax. Rates subject to change without notice Please allow up to 8 weeks for delivery of your first issue and up to 16 weeks for surface delivery outside the U S Airmail rates available on request

Payment options.
[ ] Payment enclosed [ ] Bill me
[ ] American Express [ ] MasterCard [ ] Visa

name

address

city/state/zip code/nation

card # expiration date

signature/P O #

Williams & Wilkins

P O Box 23291
Baltimore, Maryland 21203

266 Fulham Road
London SW10 9EL England

NATURE

Reader Service No 18

JOURNALS FROM MACMILLAN

☐ ANTI-CANCER DRUG DESIGN *(Quarterly)*

Editors: S. Neidle S T Crooke S. Tsukagoshi

The development of new clinically-useful anti-cancer agents is increasingly dependent on rational scientific principles drawn from a wide range of disciplines This new journal aims to promote such approaches, providing a common meeting ground for medicinal chemists, pharmacologists, biochemists, biophysicists and oncologists involved in both fundamental and clinically orientated aspects

This journal is published with the support of The Cancer Research Campaign

Reader Service No 80

☐ BONE MARROW TRANSPLANTATION

(Quarterly)

Editor: J. M. Goldman
Co-Editor: R. P. Gale

A new journal which publishes clinical results in man and experimental results with animals, allografting and autografting, HLA and other methods of selecting donors, graft-versus-host disease and all types of methods to prevent and treat it, and much more

NEW JOURNAL FIRST ISSUE PUBLISHED MAY 1986

Reader Service No 81

☐ BRITISH JOURNAL OF CANCER *(Monthly)*

The Journal of the Cancer Research Campaign

Chairman: D. G Harnden Editor: M Moore

One of the leading cancer journals in the world publishing articles on all the major scientific disciplines

Publishes papers from researchers throughout the world and the proceedings of meetings of the British Association for Cancer Research and other symposia

Reader Service No 82

☐ HUMAN TOXICOLOGY

(Bi-monthly)

Secretary to the Editorial Board: P. Turner

Covers all aspects of human toxicology embracing both animal research with direct relevance to man and studies with humans

* Now in its fifth year, *Human Toxicology* publishes all the up-to-date information in this rapidly expanding field

Reader Service No 83

N41286

SAMPLE COPY ORDER FORM

Name ____________________

Position ____________________

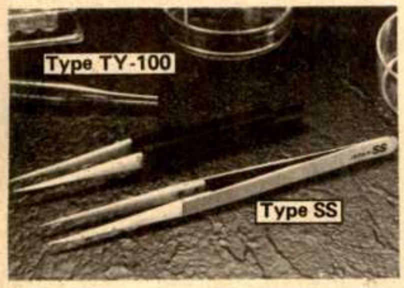
Address ____________________

____

Please tick the sample copies you would like to receive and return this advertisement to **Richard Gedye, Macmillan Press, Houndmills, Basingstoke, Hants, RG21 2XS, UK.**

Macmillan Press
Scientific & Medical

NEW CERAMICS PRECISION:
FUNA-TWEEZERS



"FUNA-TWEEZERS" can eliminate all demerits of stainless or plastic tweezers. The stainless stems of Type SS and the Derlin stems with glass reinforcing agent of Type TY-100 give appropriate elasticity.

These show excellent features such as high anti-corrosiveness, heat insulation, high flame resistance, etc. These are also free from plastic deformation, metallic contamination and magnetization, etc.

These tweezers are contributing to the industries and research such as Biotechnology, Electronics, etc.

FUNAKOSHI PHARMACEUTICAL CO., LTD.
Science Business Div. Research Material Sect.
2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan
Telephone : Tokyo 03-293-2352
Telefax : 81-3-295-5545
Telex : J28489 FUNA

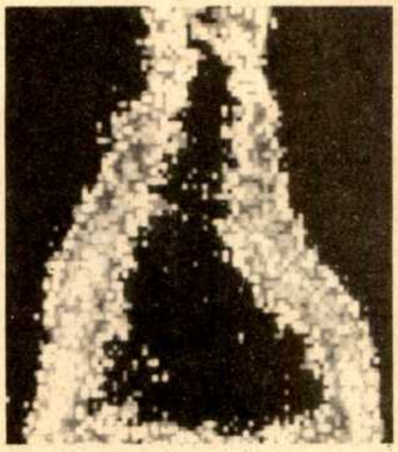
Reader Service No.73

Microtex Scientific Imaging Lets Your Computer See The Light

An integrated scientific imaging workstation from Microtex can solve your image acquisition problems, and provide results in minutes instead of days.

Whether your application is high speed, low light, UV or IR, a Microtex system can capture, analyze and display almost any image an experiment can create.

For expert applications engineering and further information, contact:



MICROTEX
DIGITAL IMAGING

80 Trowbridge Street
Cambridge, MA 02138
617-491-2874 Telex 948-014

Reader Service No.16

PLUS DE 200 NOUVEAUX PRODUITS

dont :

ANTICORPS MONOCLONAUX DIRIGES CONTRE LES ANTIGENES :

- de CELLULES HEMATOPOIETIQUES d'homme, de rat et de souris ;
- de CLASSES ET SOUS-CLASSE D'IMMUNOGLOBULINES d'homme, de rat et de souris ;
- de VIRUS, PARASITES ET BACTERIES ;
- d'HORMONES HUMAINES ET MARQUEURS TUMORAUX pour l'immunohistochimie.

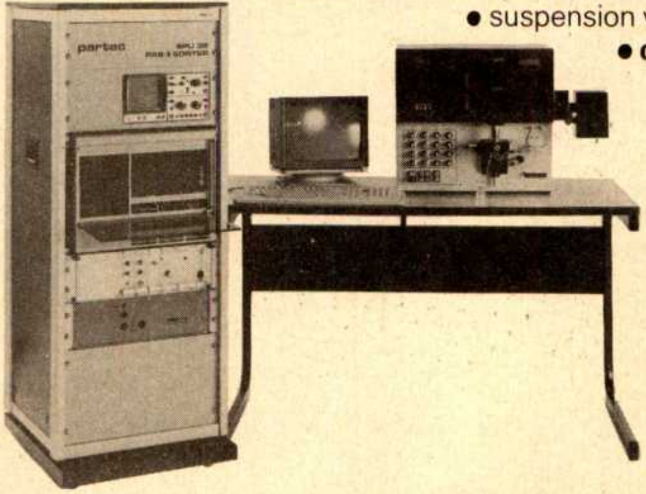
SERACLONE : RENDEMENT DE FUSION SUPERIEUR A 90 %

- Lots de sérum fœtal spécialement sélectionné pour la fusion et le clonage cellulaire sur échantillon et réservations.

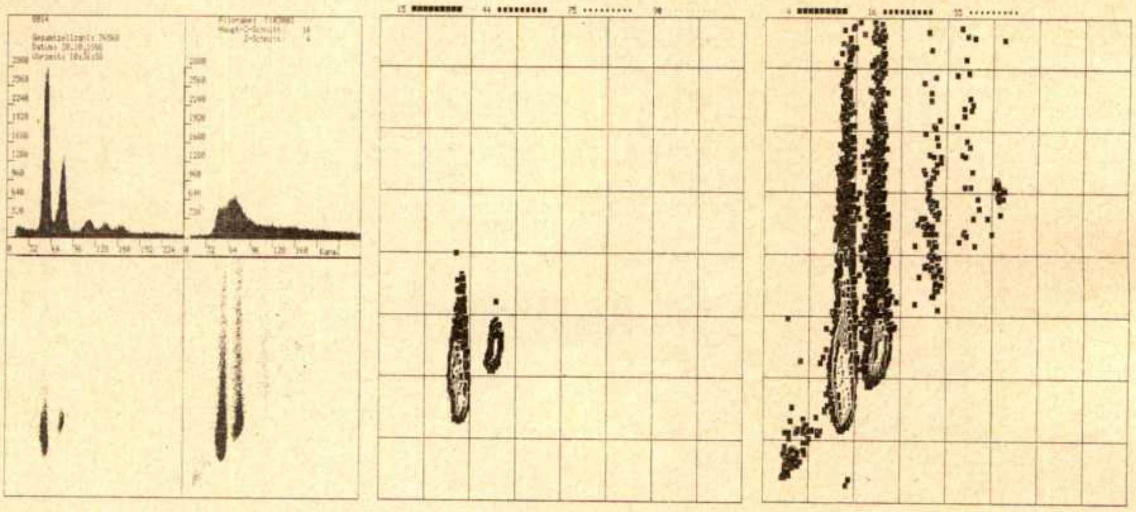
21, quai du Clos des Roses - 60200 Compiègne - France - Tél. : 44.86.22.75 - Télex : 145 171 F

Reader Service No.64

partec PAS-II THE FLOWCYTOMETER AND SORTER FOR THE DAILY ROUTINE

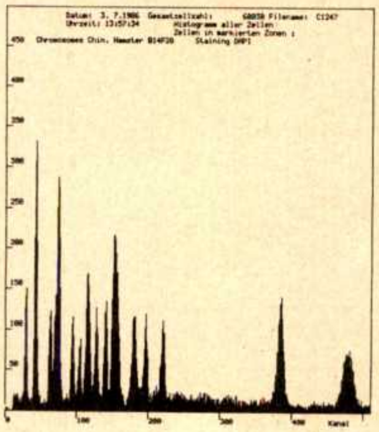


- 1–3 optical parameters (fluorescence)
- conductometric particle sizing
- suspension volume sensor for particles/ml
- **closed sorting chamber: avoids contamination of laboratory environment resp. hazard for users**
- **sorting of large, viable, and fragile cells like plant protoplasts**
- **sterile sorting of cells, chromosomes**
- 16/32 bit computer for real time acquisition and on line calculation, print out etc.
- modular system at prices between 30.000 and 100.000 US$
- 12 months full warranty



DNA-Protein Content after staining with Dapi and Sulforhodamin Human Lung-Cancer

- **Oncology**
- **Pathology**
- **Hematology**
- **Immunology**
- **Cytogenetics**
- **Microbiology**
- **Biotechnology**
- **Plant protoplast sorting**



Chromosomes-DNA-Content Distribution of CHO Chromosomes after staining with Dapi (Dr. Otto, Fraunhofer-Institut für Toxikologie, D-5948 Schmallenberg)

partec
Instruments for Research

Partec AG
Sonnenweg 7
CH–4144 Arlesheim / Switzerland
Telefon 061 72 77 55
Telex 64 740 pas ch

Reader Service No.8

Quantitative receptor autoradiography

M.D. Hall, A.P. Davenport & C.R. Clark

Quantifying receptor autoradiography is a useful tool, and one becoming increasingly popular, but potential errors exist in the interpretation of results.

RECEPTOR binding studies are powerful weapons in the neurochemist's armamentarium, their present application being predominantly in basic neurochemical research, drug discovery programmes and toxicology studies. Recently their status has been elevated by the increasing sophistication of receptor autoradiography. The predominant advantage of this technique over conventional homogenate-based binding assays is enhanced anatomical resolution combined with an increased sensitivity that is reported to be four to five orders of magnitude greater than "grind and bind" assays[1]. The additional advantage of autoradiography is that the distribution of binding sites, revealed as patterns of radioactivity, can be related to detailed structures in the specimen[1].

Here we highlight the dilemma associated with the standardization of receptor autoradiograms, and pinpoint potential errors in the associated calculations.

The technique

Receptor binding to slide-mounted tissue sections is performed essentially in the same way as homogenate-based assays. Sample sections 10 μm thick are mounted onto microscope slides, incubated with radioligand, washed and then dried. Receptor autoradiograms are produced by tightly apposing the tissue to either emulsion-coated coverslips or commercial radioisotopic-sensitive film. Following exposure and development, binding sites are visualized as silver grains in nuclear emulsions or as gray images on autoradiographic film.

Examination of autoradiographic images on emulsion-coated coverslips is facilitated by dark-field microscopy, where the silver grains appear white on a dark background. This technique gives better resolution than film autoradiography because the crystal diameter of the emulsion is smaller than that used to coat films. However, this method is capricious and quantification of silver grains is time-consuming; achieved by literally counting the number of grains per unit area. In contrast, film-based receptor autoradiography is comparatively easy to perform, and allows up to 60 microscope slides to be processed simultaneously. An example of film-based receptor autoradiography is shown in Fig. 1. In film autoradiography the gray levels associated with receptor localization can be readily converted into optical densities using an image analysis system, to allow quantification. The main disadvantage of film compared to the coverslip technique is that the tissue sections are not attached permanently to the autoradiographic image, making the identification of discretely labelled regions more difficult.

Until recently, a serious disadvantage of autoradiography compared to homogenate-based assays was that it was only quantitative. In 1981, the use of calibrated radioactive standards that were co-exposed with the tissue sections under study was introduced[4]. Two types of standards are currently used. The first is commercially available, and consists of increasing quantities of [^{3}H] or [^{14}C] embedded in blocks of polymethylmethacrylate. Alternatively, standards can be prepared[5] in the laboratory by mixing increasing quantities of radionuclide with a tissue paste prepared from the same tissues as those used in the receptor assay. For the assay, a set of calibrated standards (polymer or tissue paste) are cut at the same thickness as tissue sections and are developed alongside the tissue sections. Thus optical densities can be converted into molar quantities of receptor-bound radioligand by comparing the optical density of the tissue sample to the known standard. Figure 2 shows a photograph of a series of [^{125}I] brain paste standards exposed to Ultrofilm for 10 days (*a*) together with a natural log plot (*b*) of optical density versus radioactivity contained within each standard section.

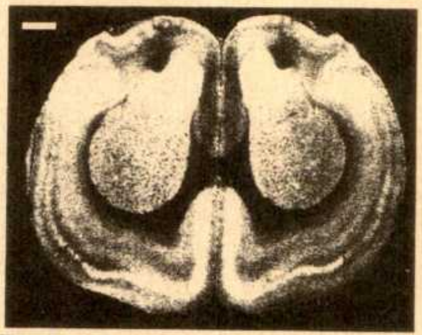


Fig. 1 *a*, An autoradiograph demonstrating film-based receptor autoradiography for the binding of cholecystokinin in rat brain.

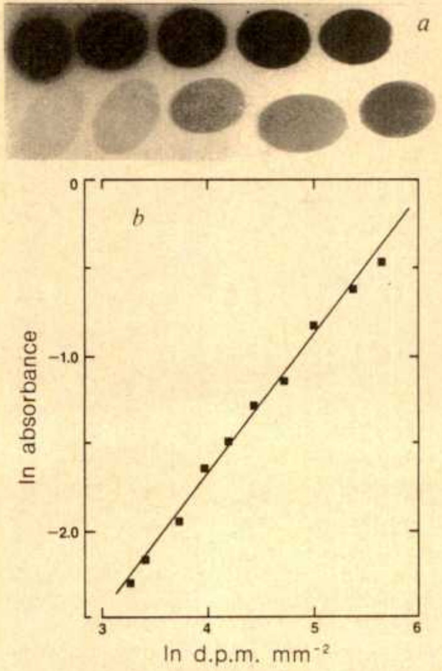


Fig. 2 *a*, Photograph of an autoradiographic image for a series of calibrated [^{125}I] brain paste standards. *b*, A natural logarithmic plot of optical density vs. radioactivity per mm^2 for the series of [^{125}I] brain paste standards[5] in *a*.

Dilemmas of standardization

A major consideration when using [^{3}H] standards is quenching, or the absorption of the low energy β-particles by the medium before they can reach the photographic emulsion. The generation of tissue paste standards is an attempt to circumvent this problem since it is reasoned that the degree of β-particle absorption in tissue paste is the same as that occurring within tissue sections. However, this does not completely overcome the problem, because the degree of quenching of [^{3}H] may differ within tissue types. Absorption in tissues from the central nervous system, for example, has been demonstrated to be different in gray and white matter[6], and methods for coping with this regional absorption have been discussed[7]. In contrast, self-absorption of β-particles from [^{125}I] has been shown in central nervous system tissues to be less significant because [^{125}I] emits electrons of higher energies than [^{3}H]. Since [^{125}I] polymer-based standards are not yet available, tissue paste standards represent the only way of quantifying receptor autoradiograms in molar terms using [^{125}I].

A second problem concerning the standardization of receptor autoradiograms using any radioisotope concerns the frequent, but incorrect interchange of two unrelated terms, "molar quantities of receptor bound ligand per mg of tissue protein" and "molar quantities per mg of standard protein". These are not synonymous, because the protein concentration of tissue paste may not be the same as that of the tissue under investigation. We therefore recommend adoption of the term dpm mm^{-2}, which can be converted into molar quantities per surface area[5] since the absolute values for amount of radioactivity, surface area and optical density can be readily obtained from standard sections.

Recently, a further problem concerning standardization has been encountered, during the conversion of dpm mm^{-2} to fmoles bound mm^{-2}. When using [^{3}H] labelled material the potential error is negligible because this isotope has a long half-life of 12.35 years. In contrast, the relatively short 60-day half-life of [^{125}I] necessitates the regular production of tissue paste standards. In practice, these [^{125}I] standards are not prepared at the same time as the binding assay, and are used repeatedly over two half-lives. Calculation errors result when the specific activity of the standard instead of the specific activity of the radiolabel used in the binding assay is used to make the conversion of dpm into fmoles. Depending upon the relative age (or decay) of tissue paste standards compared to the radiolabel used in the assay, up to two-fold under- or overestimations of specific binding can occur, and this might have contributed to the considerable differences reported between laboratories of values for fmoles of radioligand bound per unit area.

Future potential

Considerable advances have been made in the standardization and quantification of autoradiograms over the past five years. To further improve the technique, the anatomical resolution could be greatly enhanced if the crystal diameter of the emulsion used to make the films was reduced significantly. Theoretically, the sensitivity of the emulsion would also be reduced, but this could be overcome by increasing the time of apposition. The development of three-dimensional image analysis of receptor autoradiograms is inevitable with the rapidly increasing capacities of new computer systems, and this will definitely expand the range of possibilities for the applications of receptor autoradiography.

Martin Hall and Anthony Davenport are at the Parke–Davis Research Unit, Addenbrooke's Hospital Site, Hills Road, Cambridge CB2 2QB, UK, and Colin Clark is at the Pharmaceutical Research Division, Warner Lambert Co., 2800 Plymouth Road, Ann Arbor, Michigan 48105, USA.

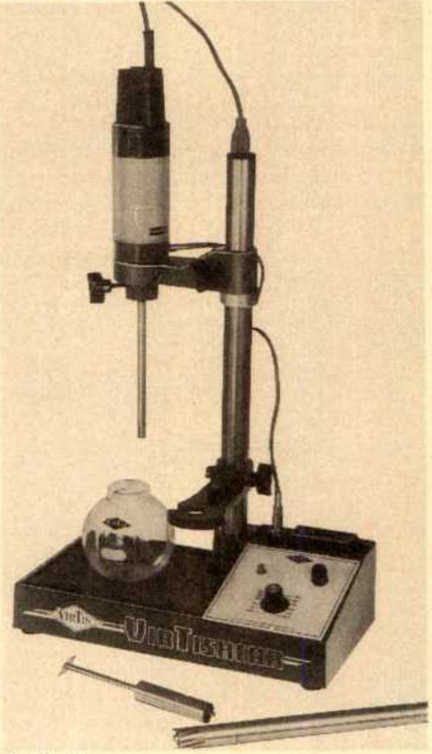
1. Rogers, A.W. *Techniques of Autoradiography*, 3rd edn (Elsevier, Amsterdam, 1979).
2. Stumpf, W.E. & Roth, L.J. *J. Histochem. Cytochem.* **14**, 274–287 (1966).
3. Young, W.S. & Kuhar, M.J. *Brain Res.* **179**, 255–270 (1979).
4. Unnerstall, J.R. *et al. J. Pharmac. exp. Ther.* **218**, 797–804 (1981).
6. Palacios, J.M. in *CVS Receptors from Molecular Pharmacology to Behaviour* (eds Mandel, P. & DeFeudis, F.V.) 455–463 (Raven, New York, 1983).
7. Kuhar, M.J. & Unnerstall, J.R. *Trends Neurosci.* **8**, 49–53 (1985).

International exhibits

Highlights from the Salon du Laboratoire and Interchemie '86 in Paris, and from the American Society for Cell Biology meeting in Washington, DC

PROMEGA'S X-Cell **electroporation systems** are designed for plant cell, mammalian, and bacterial electroporation (Reader Service No. 100). The $4,950 (US) X-Cell 450 provides a variety of parameter combinations, with a voltage range of 0–450, and a capacitance of 50–1,550 µF, used for mammalian and plant cells, while the $3,950 (US) X-Cell 2000 provides higher voltages (0–2,000 volts, 5,000 optional) and lower capacitance (14 µF; suitable for mammalian and bacterial cell electroporation. Both systems employ disposable cuvettes. Promega will be in booths 925 and 927 at the American Society for Cell Biology meeting.

VirTis now offers a **mechanical tissue homogenizer**, designed for the rapid dispersion, emulsification and homogenization of solids, biological tissue and liquids (Reader Service No. 101). The $895 (US) VirTishear can accommodate both open-blade and rotor/stator assemblies (sold separately), and all shafts can be removed for cleaning and autoclaving. The interchangeable shafts allow the operator to process samples from 0.5 ml to 2,000 ml, and standard and specialized accessory glassware is available. For reproducible results. VirTis provides an optional panel-mounted digital speed indicator. Inquire about the VirTishear and other VirTis products at booths 524 and 526 at the American Society for Cell Biology meeting.



Really churns up a culture

Controlling the temperature of samples immersed in a water bath is often critical for reaction studies. Gallenkamp is distributing a new control head by Haake that sets out to make **temperature regulation** easier (Reader Service No. 102). The Haake F3 control head monitors the temperature of the sample directly from a pt100 sensor positioned in the sample. Temperature controllers are available for a wide range of Haake baths for temperatures between — 100 °C and 250 °C, and can be coupled to accessory programmers or computers. Haake's products will be at the Salon du Laboratoire.

Alfa-Laval has been in the fermentation business for a long time, and provides equipment scaled down for the research laboratory, as well as larger setups for industry (Reader Service No. 103). In addition to fermenters, Alfa-Laval has created a **vertical stirrer**, the Vibro-Mixer. The stirrer consists of horizontal plates with vertical cone-shaped nozzles, and as the stirrer moves up and down, liquid is pumped through the cones by the venturi effect. Alfa-Laval says it is ideal for mixing protein solutions that must not be subjected to shear. The laboratory-sized 4-litre E-1 model has a sterile seal, and starts around £500. Interchemie will have an exhibit by Alfa-Laval.

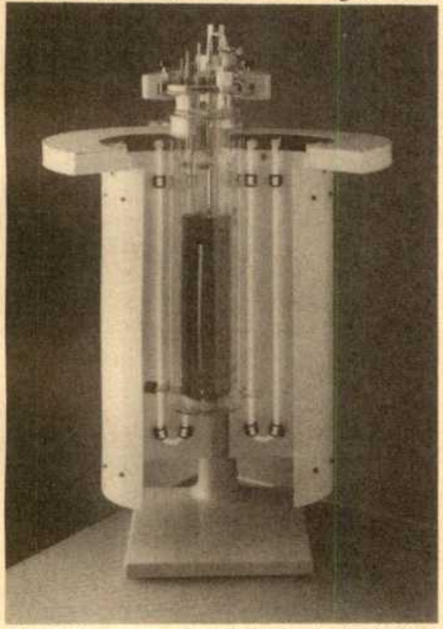
Growth enhancers

Amgen Biologicals is bringing a host of new **recombinant growth factors** to market (Reader Service No. 104). Platelet derived growth factor, the principal mitogen responsible for the growth of many serum-dependent cell cultures of mesenchymal origin, is prepared from a eukaryotic recombinant DNA host harbouring the *v-sis* gene from simian sarcoma virus. Although it differs from the cellular peptide in four amino acids, it has the same spectrum and potency of biological activity, according to Amgen. Also produced by Amgen is a potent mitogen for vascular endothelial cells and other cells of mesodermal and neuroectodermal origin, basic fibroblast growth factor, for the maintenance of cultures requiring serum-free media preparations. A newly cloned granulocyte colony stimulating factor is offered for research investigating the maturation of neutrophils and the activation of granulocyte-mediated immune responses, and tumour necrosis factor is available, shown to have tumoricidal activity *in vivo*. The growth factors are sold as 5–100 μg samples ranging in price from $225 to $775 (US). Amgen will be exhibiting at the convention of the American Society for Cell Biology in booths 19–21.

For the **radioimmunoassay of growth factors**, Amersham has launched a new line of reagent packs (Reader Service No. 105). The kits for IL-2 and IGF-1 contain one iodinated tracer, matched antisera, a buffer formulated for use in Amersham's Amerlex-M magnetic separation system and a protocol booklet. Ask about this and other Amersham products at its booth at the Salon du Laboratoire and in booth 1018 at the American Society for Cell Biology meeting.

Cell specifics

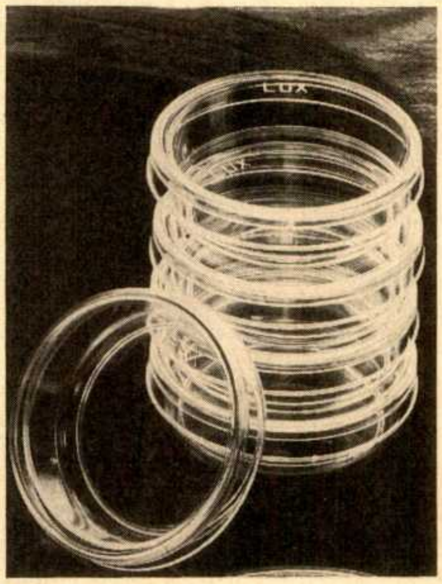
For the **cultivation of photosynthetic organisms**, LH Fermentation Ltd makes a light bank for the LH500 series of airlift fermenters (Reader Service No. 106). The light bank is mounted on the pillar of the airlift fermenter stand, and fully encircles the fermenter with light. The £1,329 (UK) light bank consists of two semicircular halves, each containing three U-shaped fluorescent lamps, that open to allow access to the fermentation vessel. The intensity of illumination can be varied to accommodate the needs of different cells. See LH Fermentation's products at the Interchemie Exhibition.



The LH Fermentation light bank sheds light upon photosynthetic cell cultures, stimulating their growth.

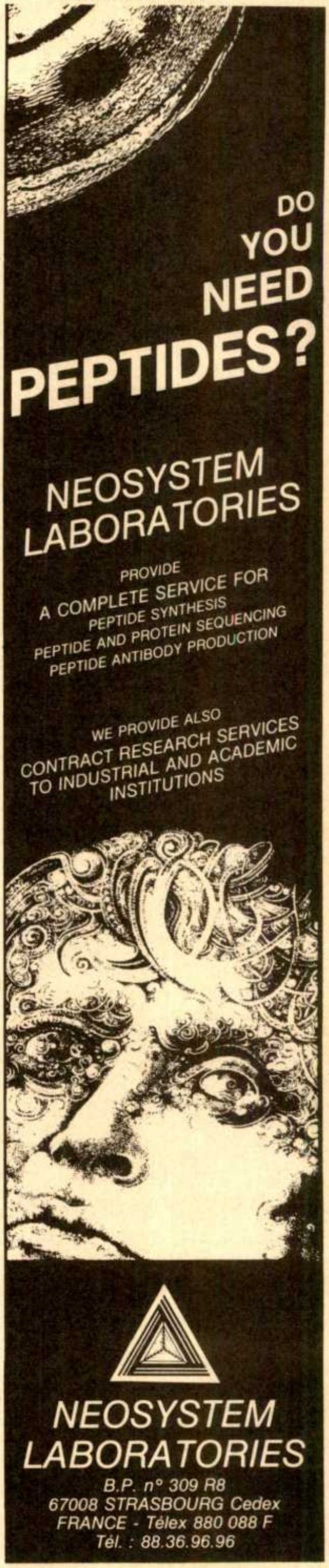
To sterilize complex culture media without autoclaving, Gelman Sciences, Ltd produces a disposable, **ready-to-use capsule filter** (Reader Service No. 107). The presterilized capsule incorporates a double layer 0.2 μm low-protein binding HT Tuffryn membrane that sterilizes up to 10 litres of serum-free media. The capsules cost £10 each. A selection of Gelman Sciences' other cell culture products will be on display in booth 514 at the American Society for Cell Biology meeting.

The Permanox **tissue culture dishes** developed by Miles Laboratories are formulated to transmit extremely low levels of molecular oxygen, thereby facilitating the anoxic growth of cancer cells for research purposes (Reader Service No. 108). The non-styrene material is also highly resistant to chemicals such as acetone and propylene oxide, so cultures may be grown, fixed and embedded for electron microscopy in the same dish. The Permanox products are a 35 × 10 mm dish, a 60 × 15 mm dish, and a 25 × 76 mm slide. The new 35 × 10 mm dishes are sold for $189 (US) in lots of 500. Miles Laboratories will be exhibiting at the Salon du Laboratoire, and at booth number 623 at the American Society for Cell Biology convention.



Grow, fix and embed for electron microscopy all in one oxygen-free dish

Uniscience, Ltd. is now distributing Wheaton's **roller culture apparatus** (Reader Service No. 109). Wheaton's modular cell production apparatus is designed for the large-scale propagation of monolayer cell cultures, and up to eight decks, holding five bottles each, may be added to the base unit. The £3,600 (UK) roller (including eight decks) features a control unit with a locking shaft, and rotates the bottles at speeds of between 0.125 and 6.9 r.p.m. with an accuracy of ±1%, according to Wheaton. See Wheaton's other wares at their booth at the Salon du Laboratoire, and at booths 610 and 612 at the American Society for Cell Biology meeting.



ADVERTISEMENT

Roland GRUNBERG · NANCY · FRANCE

DO YOU NEED PEPTIDES?

NEOSYSTEM LABORATORIES

PROVIDE
A COMPLETE SERVICE FOR
PEPTIDE SYNTHESIS
PEPTIDE AND PROTEIN SEQUENCING
PEPTIDE ANTIBODY PRODUCTION

WE PROVIDE ALSO
CONTRACT RESEARCH SERVICES
TO INDUSTRIAL AND ACADEMIC
INSTITUTIONS

NEOSYSTEM
LABORATORIES
B.P. n° 309 R8
67008 STRASBOURG Cedex
FRANCE - Télex 880 088 F
Tél. : 88.36.96.96

Reader Service No.15

ADVERTISEMENTS

LIPOSOMES

NEW

REPRODUCIBLE – HOMOGENEOUS – EXTREMELY RAPID

Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANORM

POB 126, D-8000 Munich 65, FRG

Reader Service No.47

LIPID PRODUCTS

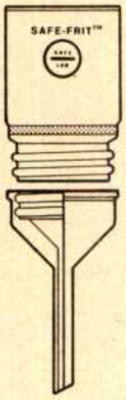
top quality lipids for biomedical research

17 years experience supplying lipids to top research laboratories throughout the world

For free catalogue contact:
Lipid Products, Nutfield Nurseries
South Nutfield, Redhill
Surrey RH1 5PG, UK
Tel: (073 782) 3277

Reader Service No.76

SAFE-FRIT™ FUNNELS



Safe-Lab's NEW fritted glass funnels feature a removable, unbreakable polypropylene stem just below the fritted glass disc. This unique design eliminates the danger of broken glass funnel tips and permits easy cleaning from both sides. Removable stem allows you to set the funnel top on counter or directly inside an oven or desiccator for drying, without the danger of tipping or spilling. Polypropylene has excellent chemical resistance and threads will not leak, even at high-vacuum.

Call or write for a FREE catalog: Safe-Lab, Inc., P.O. Box 1290, Santee, CA 92071 (619)448-9100

Reader Service No.28

Medwire Div.
stocks a wide variety of
TEFLON COATED*
Platinum alloys, stainless steel, silver, multistrand and single strand available for prompt shipment. Write for descriptive brochure and price lists.

Medwire Div.
Affiliate of Sigmund Cohn Corp
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300

*Reg. T.M. of E.I. Dupont

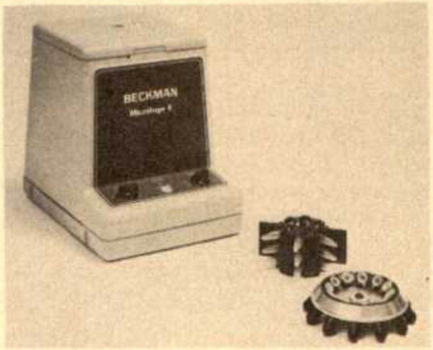
Reader Service No.49

Sample spins

The CT411 **thermostated benchtop centrifuge** sold by Jouan, Inc. has both a refrigeration and a heating unit, with the temperature range 0–60°C (Reader Service No. 110). Jouan says its $5,500 (US) centrifuge can attain a speed of up to 5,840*g* and that its digital readout is accurate to 10 r.p.m. Standard features are a safety lid interlock, a continuously variable braking rate and an imbalance detection and shutdown system. Jouan will have a booth at Salon du Laboratoire.

Spins hot or cold

Beckman sells a miniature **tabletop centrifuge** that it says is ideal for repetitive runs of small biological samples (Reader Service No. 111). The $845 (US) Microfuge E uses either a horizontal rotor ($130 US) capable of speeds of up to 14,500 r.p.m. at forces of up to 13,490*g*, or a fixed angle rotor ($150 US) that runs up to 15,000 rpm at centrifugal forces of up to 15,850*g*. It holds from 12 to 48 microtubes, 250 µl to 1.8 ml in size and has a "momentary" button for one-second spins. Beckman will be displaying its centrifuges at the Salon du Laboratoire and at the American Society for Cell Biology meeting.



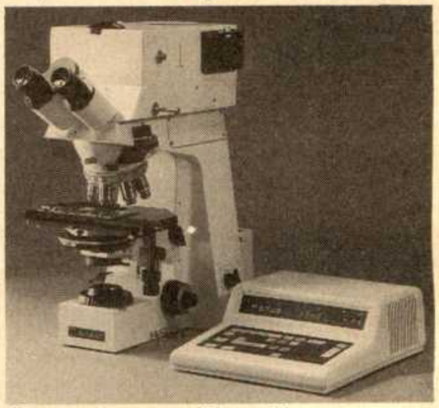
Just the thing for tiny amounts

To make sure the right bottle is being used for the right rotor, Nalgene has compiled a *Centrifuge Tube and Bottle Rotor Matching Guide* (Reader Service No. 112). The free 12-page guide contains tables matching the appropriate Nalgene tube or bottle to the rotors of nine different manufacturers. Tips on sterilization, sealing caps and cap assemblies, and a chemical resistance chart are also included. Nalge has an exhibit at the American Society for Cell Biology meeting, booth 301, and will also be present at Salon du Laboratoire.

Costar's **microcentrifuge filters** can filter volumes of 25 µl to 500 µl with a retained volume of less than 5 µl, according to the manufacturer (Reader Service No. 113). The Spin-X filters are available in two pore sizes, 0.22 µm or 0.45 µm, and several membrane types, including an ultra-low binding cellulose acetate. Costar is in booths 54 and 536 at the American Society for Cell Biology meeting.

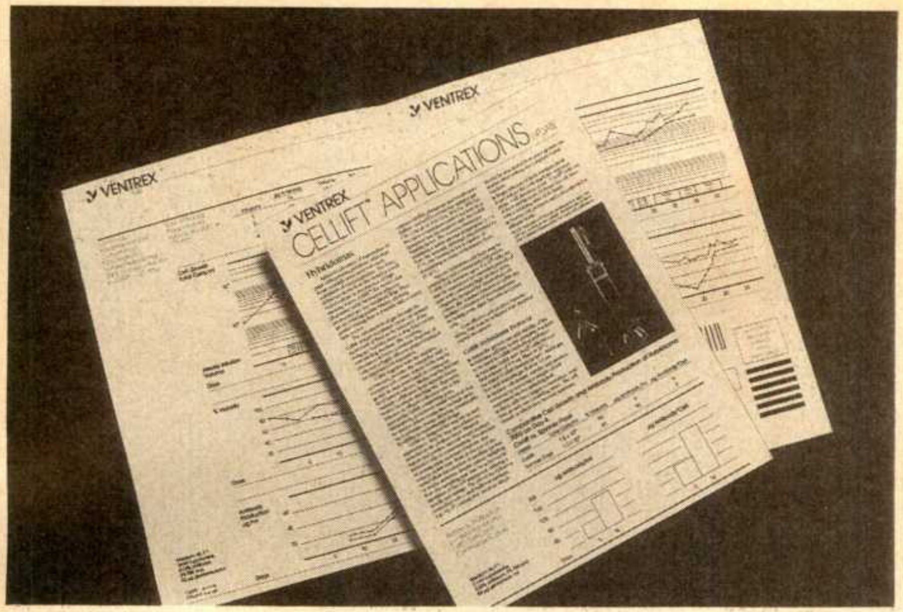
The better to see...

For improved resolution and image quality in **photomicroscopy** with no sacrifice in field of view, Carl Zeiss, Inc. has developed the Axiophot advanced research microscope (Reader Service No. 114). The Axiophot contains ICS (infinity-colour-corrected system) optics to ensure clear, aberration-free images over a 25 mm field, according to Zeiss, and all components are system-integrated (SI Design). The Axiophot has two 35 mm built-in microprocessor-controlled cameras operated with an external keyboard, and carries a $32,000 (US) price tag. Carl Zeiss has an exhibit booth at the Salon du Laboratoire and will be in booths 618–626 at the American Society for Cell Biology meeting.



Captures important slides on film

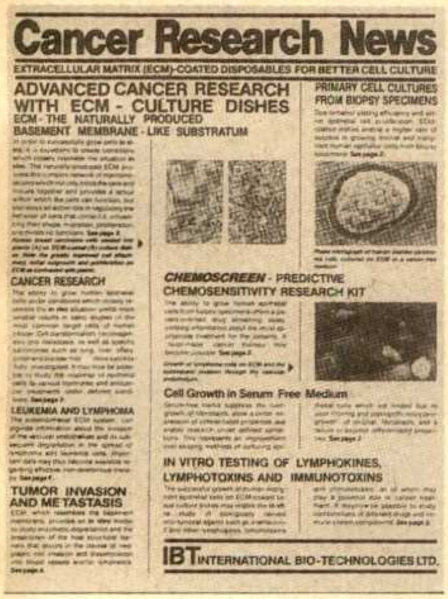
The PL Fluotar series of **microscope objectives** from Leitz are ideal for fluorescence microscope and interference contrast work, according to the manufacturer (Reader Service No. 115). They offer high near-ultraviolet transmission, minimal autofluorescence and negligible strain, and provide neutral colour reproduction and a field flat to 25 mm. The objectives range in price from £329.50 (UK) for the 6.3× objective, to £348 (UK) for the 25× objective. More information on Leitz's objectives will be available from the Leitz booth at the Salon du Laboratoire, and from booths 1036 and 1038 at the American Society for Cell Biology meeting.



The book with the lowdown on the Ventrex range of airlift fermenters

Lots of literature

A recent release from International Bio-Technologies Ltd, ***Cancer Research News***, outlines the role of extracellular matrix (ECM) in the field of cancer research (Reader Service No. 116). The six-page brochure aims to introduce International Bio-Technologies' line of ECM-coated labware, but also contains several interesting articles on the culture of cancer cells in general. *Cancer Research News* also offers an introduction to International Bio-Technologies' library of other free brochures, on topics ranging from immunology to neurobiology. International Bio-Technologies will have a display in booth 1035 at the American Society for Cell Biology meeting.


Cancer Research News

EXTRACELLULAR MATRIX (ECM)-COATED DISPOSABLES FOR BETTER CELL CULTURE

ADVANCED CANCER RESEARCH WITH ECM - CULTURE DISHES

ECM - THE NATURALLY PRODUCED BASEMENT MEMBRANE - LIKE SUBSTRATUM

PRIMARY CELL CULTURES FROM BIOPSY SPECIMENS

CANCER RESEARCH

CHEMOSCREEN - PREDICTIVE CHEMOSENSITIVITY RESEARCH KIT

LEUKEMIA AND LYMPHOMA

Cell Growth in Serum Free Medium

IN VITRO TESTING OF LYMPHOKINES, LYMPHOTOXINS AND IMMUNOTOXINS

TUMOR INVASION AND METASTASIS

IBT INTERNATIONAL BIO-TECHNOLOGIES LTD.

International Bio-Technologies highlights the latest techniques in culturing cancer cells

Ventrex Laboratories, Inc. will be on hand in booth 806 at the American Society for Cell Biology meeting, and will be able to provide information about ***Cellift Applications Update*** (Reader Service No. 117). This technical publication outlines the procedures for using Ventrex's Cellift disposable airlift bioreactor. Ventrex says Cellift's advantage is the provision of a well-mixed culture in the absence of high shearing forces, and is ideal for the culture of a wide range of cells, from mammalian to yeast.

Free **custom culture optimization** is only one of the services provided by Scott Laboratories (Reader Service No. 118). With the guaranteed purchase of a minimum order of growth media, Scott will work out the optimum growth conditions of a customer's culture, using a client-supplied cell line sample and analysis of spent media. Scott claims a 70 per cent success rate for its service so far, with past results yielding 2–50 fold increases in product. Scott's new catalogue, giving full descriptions of their other services, media, and growth factors, will be on display at booth 330 at the American Society for Cell Biology meeting.

An easy **wall chart reference for the preparation of cell culture media** is available free from Imperial Laboratories, as a product insert (Reader Service No. 119). It gives the formulations of the most commonly used media, including information on reconstitution, and quantities of sodium bicarbonate and L-glutamine required for various media. ☐

These notes are compiled by Carol Ezzell from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated.

ADVERTISEMENTS

LIMUTESTER

(gelation-capillary method)

Precise Detection of Endotoxins

This LIMUTESTER detects Endotoxins promptly, applying to gelation reaction of LAL and a glass capillary.
The LIMUTESTER is simple, economical and usable anywhere, especially needs no endotoxin-free laboratory glass wares.
LIMUTESTER INCUBATOR is available.

FUNAKOSHI PHARMACEUTICAL CO., LTD.
Science Business Div. Research Material Sect.
2-3 Surugadai, Kanda, Chiyoda-ku, Tokyo, Japan
Telephone : Tokyo 03-293-2352
Telefax : 81-3-295-5545
Telex : J28489 FUNA

Reader Service No.75

WESTERN BLOTS IN MINUTES

Our new GENIE electrophoretic blotter incorporates space-age plate electrodes for a uniform, high-intensity field. It is easy to load, and uses less than a liter of buffer, yet can blot gels up to 15 by 17 cm. in as little as ten minutes.

only $190

Idea Scientific Company
P.O. Box 2078
Corvallis, Oregon 97339
(503) 758-0999

ASCB Booth No. 326

Reader Service No.14

Versatile Tissue Culture Incubator

Modular Incubator Chamber (the Cell/Sphere™) for all tissue cultures, aerobic or anaerobic.

Low Cost

Eliminates evaporation loss and contamination. Easy to use – flush with gas mixture, seal, place at appropriate temperature. Stackable, see-through. Quantity discounts. **billups-rothenberg, inc., pob 977, Del Mar, CA USA 92014. Call 619-755-3309 or contact your Flow Laboratories representative.**

Reader Service No.7

Strategy for higher education

Richard Pearson

Modest increases in government spending on universities announced recently do nothing to fulfil the urgent need for a long-term strategy for higher education.

It is now 18 months since the government published its consultative Green Paper on the future of higher education. Since then higher education has staggered on, responding to piecemeal evolution of a policy, which until last month has mainly resulted in a tightening of the financial screw. November saw a slight improvement, however, with a small increase in the universities' budgets as part of the government's boost to public spending. Whether this signals a change of policy is not clear and the government's White Paper on the future of higher education still seems many months away. Higher education will remain in a state of crisis and policy decisions by students and institutions will have to be taken on an *ad hoc* basis until the government sets out a long-term strategy.

Although some academics still fear a threat to academic freedom, it is clearly right that the role and running of a higher education system costing the public purse over £3,500 million each year should be a matter for public debate and scrutiny. But it would be wrong to seek change through a series of partial solutions and short-term actions that cause damage in the long term, while providing no strategic direction for the future. Major unresolved issues include those of student numbers, subject balance, student finance, and academic salaries and resources.

In the past the driving force influencing the size of higher education has been student demand, and particularly the wishes of 18-year-olds. But, with a student population of about 500,000, the United Kingdom has low participation rates in higher education by international standards especially by women in technological subjects, mature entrants and ethnic minorities. The proportion of entrants from working class homes, at under 10 per cent, has barely changed for a decade. Participation by these groups needs to be increased, both for social reasons and to broaden the quality and range of experience in the graduate population which is too narrowly based. Such an increase would counteract the reduced demand for places due to the expected fall in the number of 18-year-olds in the period to 1995. Demand for graduates is now at record levels and expected to increase significantly, and the longer term need is for a better educated and more skilled workforce.

These social and economic objectives of higher education need not be incompatible. For example, the labour market too needs higher education to have a broader base than is provided by simple reliance on vocational subjects, because employers need a broad cross-section of skilled and educated staff. In the vast majority of jobs and careers, indeed there is no close link with degree subject, so that one in three of current vacancies for graduates do not specify a degree subject. This allows flexibility in the provision of courses for the majority of places and responsiveness to student demand and other criteria when planning subject balance. It also allows rebuttal of the complaint that precise manpower planning does not work, because such accuracy is simply not needed.

There is, though, a need to improve the content and quality of many courses to encourage communication skills, team work, numeracy and literacy. In addition a clarification of which are truly vocational courses needs to be undertaken and a better understanding of employers' current and future needs for specialist trained graduates is required.

More students, better salaries to retain and attract good teachers and adequate, up-to-date resources all cost money. They cannot realistically be paid for by savings from a loan scheme to replace student grants. Industrial support by sponsoring students, donating equipment, funding research, paying for visiting lecturers and financing higher academic salaries, all help higher education as well as the recruitment and research of the firms themselves. There are no exact figures on the full extent of direct industrial support, but overall it is probably less than £100 million, or under 3 per cent of the higher education budget. Thus even the expected significant increases over the next decade will not be enough to relieve the burden on the public purse, though it will continue to make a valuable contribution in selected areas. While improvements in efficiency can also undoubtedly be achieved, the necessary improvements in higher education cannot be realized within a reducing budget; public financing has to be increased in real terms, not reduced.

A better way of distributing finite resources may be through the introduction of the two-year general degree, topped up by a two-year specialist degree. The first degree would be available to all; it would offer a broad-based curriculum and would attract a maintenance grant. Student demand could strongly influence subject balance as long as there were some core general subjects which would compensate for any over specialization in the schools. Such degree courses could offer far more flexible entry qualifications, which would allow more mature students and those from less conventional backgrounds to enter higher education.

There would then be a smaller number of places on second degrees which would be more specialized, focusing on particular disciplines. The balance would be determined more by national or occupational criteria, as already happens in subjects such as medicine and architecture, or by the need for advanced study, as is the case for postgraduate courses in the arts and many of the social sciences. These subjects would attract further grants, scholarships or industrial sponsorship in key areas of the labour market and loans could also be available. A critical factor in their implementation would be that all degree courses should follow this pattern — otherwise the new two-year degree courses would be instantly dubbed second class.

Whatever the strategy for higher education — and the options are diverse — there must be a long-term perspective and an improved planning framework, which would require better information about inputs, processes and outcomes. If market forces are to play a key role then participants have to be better informed: students about job prospects to help in the choice of subject and course, institutions about economic needs to tailor a proportion of the courses accordingly, and government in terms of deciding how much money to spend, and where and how to allocate it. All these groups require a better articulation of the longer-term relationship between economic and industrial needs and the provision of higher education, so that priorities can be considered, and flexibility built into the system to cope with uncertainty. The performance of higher education as a whole also requires regular monitoring and policies must respond to changing economic and social needs.

Whatever the complexion of governments in the next decade there is unlikely to be a significant increase in the funding of higher education. Piecemeal policy adjustments are likely only to reinforce the downwards spiral and inhibit a successful transition into the 1990s. A longer-term strategy is vital. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton BN1 9RF, UK.

nature

Another nuclear power station?

Not much has happened in the nuclear power industry since Chernobyl. But now the British government has to make a decision that will determine not only British plans, but influence what others do.

For understandable if not particularly good reasons, the managers of nuclear power enterprises outside the Soviet Union have been circumspectly unostentatious since the Chernobyl accident. The world's two hundred or so power reactors have kept on producing electricity, and the economic benefit that goes with it, without further incident, but otherwise it has been a slack time. Even in France, which manufacturers more than 70 per cent of its electricity from uranium, there promises now to be something of a pause in the building programme. In West Germany, Chernobyl has helped to bring nearer the end of the high-temperature reactor project at Kalkar, while talk of other fresh initiatives has almost ceased.

Part of the explanation is economic; oil prices (in absolute value) are lower than for two decades, while fears of physical shortage of oil are belied by the many decisions now being made not to exploit known reserves, from Alaska to the China Sea. But a more interesting part of the explanation is the nervousness induced by the Chernobyl accident. If a crew of foolhardy reactor operators can, by wayward disregard for operating procedures, cause an accident which, tragic even in the scale of the Soviet Union, could have made many smaller countries uninhabitable, can any reactor be called "intrinsically safe"?

That is the dilemma that now confronts Mr Peter Walker, Secretary of State for Energy in the British government, who last week took delivery, at last, of the report of the government inspector, Sir Frank Layfield, who has been inquiring for the past three years into a proposal by the Central Electricity Generating Board (CEGB), the still-nationalized British electricity utility, to build a pressurized-water power station at Sizewell on the Suffolk coast. Walker, and the British government as a whole, may be forgiven for thinking it bad luck that they must decide on Sizewell now, less than eight months after Chernobyl and, even worse, no more than 18 months before the next general election in Britain.

So it is. The Layfield process has been under way, on one view, since 1972, which is when CEGB told a parliamentary committee that it would like to give the pressurized-water reactor a try. Five years later, Mr Tony Benn, one of Walker's predecessors in a previous government, said that he would sanction such a development only after a public inquiry. When in 1979 Mrs Margaret Thatcher, newly elected as prime minister, outlined a programme to order roughly one nuclear power station a year for the then decade ahead, she seemed not to feel bound by Benn's commitment. But in the event, British reactor designers have been occupied since 1981 in preparing material for Layfield. The long process could have ended at many other more convenient times. The only consolation, for Walker, is that somebody must jump first.

Content

In the British manner, the Layfield report has not been published. Walker would be technically within his rights to decide that it should never see the light of day, although that is so obviously a recipe for ructions as to be unthinkable. In any case, the content of the document, and its faults, are easily surmized. To the extent that safety of nuclear reactors in general and of pressurized water reactors in particular were central to the inquiry, the inspector can only say that there is no essential reason why they cannot be operated safely; to demonstrate that he was not asleep during the two long years of the public hearings, he is bound to seize on a few points in the design at which assurance could be made doubly sure by further attention to safety measures. If it had not been for Chernobyl, CEGB and its design partner Westinghouse would have sighed with relief, budgeted for the extra costs and ordered the equipment. The most obvious point on which the report (when eventually published) will be attacked is that the public hearings were completed a year ago, before Chernobyl. If the inspector has been prudent, he will have added an appendix showing that he has thought of how the inquiry might have been changed by that event. If he has not, Walker (for everybody's sake) will no doubt ask him for another piece of paper.

Scope

As public inquiries go, the Layfield exercise was unusual not only because of its length but for its scope. The intention had been to settle once and for all the issue of whether pressurized-water reactors can be safely operated in Britain, so that it might be possible to avoid the same procedure whenever another of the type seemed desirable. But, in the event, the proceedings took in all the questions relative to the good sense of generating electricity from uranium. Is all that electricity needed anyway? Will not the continued manufacture of plutonium assist the spread of nuclear weapons? And what will be done with the radioactive waste? The inquiry might have been a unique exercise in public education if it had been joined by a representative section of unofficial opinion-holders (which would have been more likely if half the time had not been spent in a remote seaside town). A wise inspector will have said that these general matters are not for him but for others, mostly for the government, the British parliament and its electors, and will then have gone on to make a few privileged unassailable observations on them. Alternatively, he may have taken a view on some general issue of policy whose delegation to him would be constitutionally improper, and may have undermined the value of his report as a whole. Either way, accepting that the British need be no less successful at operating pressurized-water reactors than people in a dozen other countries, the hard question will be left with Peter Walker.

Is nuclear power necessary, for example? That question, disarming as it may seem, is too simply asked. In the economics of energy supply, there are no categorical imperatives. Oil is an important source of energy in most industrial economies, and is a particularly convenient source of liquid fuels and of hydrocarbons in general, but economies denied access to petroleum could probably manage in other ways, but at a cost. The question to ask about the usefulness of nuclear power stations is one about the economic benefits that may be derived from them, which can be answered fully in one place only with a knowledge of what is likely to be the pattern of their use elsewhere. It goes without saying that, in present circumstances and especially when oil is cheap, it could well emerge that extra safety measures could price even the most efficient reactor designs out of the market. But Walker will also know that the fluctuating price of oil can only, in the long run, increase relative to the prices of other fuels. The need for nuclear power in Britain or anywhere is not an absolute but a matter for intricate judgement. Walker's decision will be the more convincing if he makes that plain. The Layfield report itself could turn out to be an invaluable guide to the often contradictory economics of those with opinions of the value of nuclear power — optimistic, sceptical and downright perverse.

Research

For the research community, another economic question may be even more important. Britain is not the only country whose government has decided that the technology of nuclear power is potentially too important to be left to others. Direct spending of taxpayers' money on research and development in nuclear power is probably not much short of £100 million a year, with perhaps a comparable sum creamed off electricity bills. If the decision goes against Sizewell, it will make little sense to keep up the research. But a decision to build the Sizewell plant will leave unanswered the question how research and development in this field should be financed in relation to the numbers of nuclear power stations likely to be built in the years ahead. Walker or one of his successors is going to have to tackle that question soon. It should be more easily decided now that the UK Atomic Energy Authority has been made a trading fund, required to keep books as if it were a company.

The more immediate question of safety is similarly not absolute. But the fact that there is not yet an absolutely safe nuclear reactor does not imply that reactors cannot safely be accomodated with people in the same tract of land. Both Chernobyl and, earlier, the Three Mile Island reactor accident in the United States, were illustrations of how reasonably (but not perfectly) designed machines were induced to behave badly by their operators. The Three Mile Island accident is the more worrying because the operators acted sensibly in the period leading up to the accident. The moral, for those who would build reactors, is that much more attention even than at present must be lavished on the training, discipline and performance of the operators. Those in charge must be regarded as if they were airline pilots; accidents develop more slowly in nuclear reactors than in aircraft, but their consequences are greater. If Walker wishes to convince the British public that the Sizewell station could be safely built, he could do worse than to offer what countries such as Britain need in any case (and which is partly provided in the United States by the cumbersome procedures of the Nuclear Regulatory Commission), regular and public invigilation of the management of these plants.

But what should the minister and the British government decide? The Sizewell inquiry may have dealt with the need in Britain for a nuclear power programme as such, but the immediate proposal is that there should be a single pressurized-water reactor station five years or so from now. A single plant, especially one built in the present climate, is almost certain to be put together accurately and well. CEGB says it wants to build the machine; if it can confirm, in the light of information gathered since the Layfield inquiry ended, that the electricity is still worth having at the price, why not let it go ahead? The fact that the Labour Party has said that, if elected at the general election, it would "phase out" nuclear power, in any case argues for such a step-by-step approach. No doubt a decision by the government to sanction a single plant would be called provocative, but this is part of a legitimate disagreement; but committing public resources now to a long-term programme could risk an imprudent waste. What nuclear power in Britain and elsewhere now most needs is a show of cautious deliberation by sponsors willing to insist that its virtues are judged on strictly economic grounds and that its custodians, the plant operators and their bosses, shoulder their public responsibilities squarely. □

Shades of 1968?

The Chirac government in France seems to have made a hash of its intended university reform.

JUST why General de Gaulle's government took such fright when the students built barricades in the streets of Paris in 1968 has never been crystal clear. The immediate fear was that the survival of a nationalist government representing a deeply conservative people would itself be threatened if the students' anarchical protest at a modest package of educational reforms were allowed to run its course. But the whole of France was shocked by the violence of the students' anger at M. Edgar Faure's proposals for reforming the French universities. The police fought the students at the barricades, but the government withdrew its reform package, replacing it by a more anodyne reform.

It is odd that the new Chirac government should have fallen into the same trap. Not much seems to have changed in a decade and a half. Although the Chirac government's proposals for university reform have many objectives in common with those of 1968, the ostentatious orderliness with which last week's protest began shows that the students were not at first out to make a revolution. The death of one student at the weekend after a beating from the police has inevitably changed the mood. The government, having backed down half-way in stages, is likely now to have to give in altogether.

So are the reforms dead and buried? Not quite. The issues that have angered the students will surface once again in the next few years. So they should, for they are necessary reforms. The complaint against the Chirac government is not that it has made itself look weak and even foolish, nor that it launched the reforms in the first place, but that, with such an important objective to accomplish, it assumed that it would be sufficient to use its majority in the French parliament to get the reforms on the statute books, without attempting to carry academic opinion with it.

The truth, nevertheless, is that the French university system is one of the more enduring relics of the eighteenth-century revolution. The principles are grand, as has been their practice for much of the intervening period. Able teenagers at generally excellent secondary schools qualify for university entrance by means of the nationally uniform school-leaving examination and are then, in principle, free to choose the university at which they enrol. France being France, it is inevitable that the University of Paris should have had to grow fastest under the pressure of unregulated demand, but most good institutions have been overwhelmed by numbers. One of the government's now-postponed goals was that institutions should be more able to regulate external demand by setting their own entrance requirements; a by-product would have been an element of diversity within a system which, for all its virtues, remains too much of the same piece. Can that be bad?

So what went wrong? The French government's most obvious miscalculation is to have underestimated the passions attaching to the right to university education. But in the absence of positive proposals for helping students find and choose the institutions that would best meet their individual needs — financial assistance for students in France goes very little further than the nominal tuition fees (which were in any case to have been increased) — it is inevitable that the proposals would have seemed unjust to many potential students. That is why whichever minister of education brings these proposals back to the French parliament should take care to do so within a wider framework. France is no different from most other countries in Western Europe in facing a long period during which the student age-groups will decline numerically just at the time when greater skill and technical competence are more necessary than ever. The case for managing an element of diversity within the university system is strong, but not at the cost of the quality and even quantity of the students it trains. □

After the riots the politics

Chirac drops university reform bill to buy time

FRENCH Prime Minister Jacques Chirac appeared on national television on Monday morning to save his government. He announced the abandonment of the university reform bill that had led to demonstrations, then riots and finally the death of a student in Paris on Friday evening. Chirac's fragile coalition of right-wing parties was in danger of breaking up under the pressure of these events. Alain Devaquet, the universities and research minister, handed in his resignation on Friday night after the education minister René Monory announced, also on television, that he was to take over the steering of the bill; on Sunday, Alain Madelin, industry minister, said that the reforms "were not worth a life"; and others had begun to distance themselves from an administration that was rapidly being tainted as "anti-youth".

Just eight days earlier, Chirac had offered concessions to students on the three key points at issue — selection on university entry, local university degrees (they are at present national) and an increase in entry fees, which would vary from university to university But these proved not to be enough for student leaders who demanded complete renunciation of the whole bill.

Some, though, believe that the battle was won not by the students — who claimed to be apolitical in their protest — but by more violent elements that began to attach themselves to the previously peaceful demonstrations.

Where this debacle leaves the French universities and research is by no means clear. Devaquet had spoken of the "irony" of being tarred by students as an extremist, when his whole effort has been to tone down some of the more extreme libertarian elements of the bill that emerged from the right wing of Chirac's party and while he has successfully defended against that same wing other academic institutions such as the national research centre (CNRS). His resignation leaves a gap which Chirac will find difficult to fill.

But is Chirac's attempt at university reform dead?. Chirac is an authoritarian (if realistic) prime minister, who said only a few days ago that he had been elected to "govern", so a new bill may yet emerge once tempers are cooled.

As to why the French students — and high-school children — were so determined to resist selection, and increasing university autonomy, the answers are more social and psychological than political.

5 December: a demonstrator with sling faces riot police.

Selection does operate in French universities already, but it bites late on entry to the second and third years. It is made by university staff on new grounds; so a schoolboy or girl who has failed to make good progress at the very authoritarian French schools has a second chance in the first year of university (provided he or she passes the school-leaving *baccalaureat*). Pre-university selection means denying that second chance; and that can mean the dole queue.

The previous government planned to increase the entry to the first year of university, and making it explicitly a year of 'orientation' that could lead to jobs or other educational institutions as well as onwards in the university; but it failed to provide the resources to accomplish the task and created resentment among university staff. This government attempted to please that staff. But it forgot about the students. **Robert Walgate**

Vera Rich was in Paris when the students' patience ran out . . .

PARTICIPANTS in last week's clashes between French university students and the police were at a loss to explain how matters had got so out of hand. What was planned as a peaceful, silent, non-political protest on Thursday escalated into disturbances which, on Friday night, culminated in the death of the 22-year-old student Malik Oussekine. Certainly the Thursday demonstration was "politicized" by participants who, contrary to the organizers' appeals, brought political banners. But a high police profile at the march aggravated the situation. By Friday, the nationwide protests had taken on a frankly political tone, and on Saturday trades union leaders were being asked to sit in on the meetings of student delegates from all over France, and calls were going out for a one-day general strike.

The government's reactions — as so often in such situations — were too little, too late. The first, partial climb-down over the bill might have been effective on Thursday, before violence developed but it came only on Friday, a few hours before Oussekine's death. Security Minister Robert Pandraud's statement, on Saturday, that if there had been "individual faults" on the part of the police, sanctions would be imposed, was coupled with a refusal to dissolve the special police 'anti-vandal' motorcycle squadds introduced early on Saturday morning.

Interestingly, by Saturday night, even students involved in the growing disorders were convinced that the genuine protests had been "manipulated", although opinions differed as to whether the alleged manipulators came from the extreme right or the extreme left.

From the beginning, there has been an emotional and in some degree irrational element in the students' and *lycéens'* opposition to what, viewed more logically, might have been seen as an unpleasant but necessary reform, given the current economic state of France and the problems of graduate unemployment. But *lycée* scholars have grown up considering it an inalienable right that all school leavers who pass the *baccalaureat* can automatically go to university, while restrictions on choice and combinations of subjects, and the possible development of a hierarchy of excellence among the universities, seemed a direct affront to liberty, equality and fraternity. Some students, too, seem to have been stung into their decision to take part in the demonstrations by the accusation, voiced on French television only a couple of weeks before by Daniel Cohn-Bendit, the 1968 student leader, that the present generation of French students are "apolitical".

In reality, as several university lecturers have been at pains to point out, the hierarchical principle has long been established *de facto* in France, with prospective employers at least as interested in where an applicant studied as in his or her subject and academic record. The new law, they said, would simply validate the existing system. But as early as Friday afternoon, and particularly since the death of Oussekine, these same lecturers began privately to adopt a different stance. "I do not know whether I am on the side of the students over the law", one said, "but I am certainly not on the side of the police against the students!" Vera Rich

European summit

Unanimity but little of substance

EUROPEAN leaders confronted with the massive overproduction and costs of the Common Agricultural Policy and the need for technological cooperation to face competition from the United States and Japan, laboured mightily to produce a mouse last week: an anodyne communiqué, the result of the London summit marking the last month of Britain's six-month presidency of the European Communities, which failed to mention agriculture at all, put off budget questions to a later summit, and included lukewarm support for research in Europe only on the last-minute insistence of the French President, François Mitterand.

This time, the presidents, prime ministers and chancellors of the 12 states of the European Communities were in no mood to tackle controversial topics, at least in public. Foreign Office sources in London indicated not entirely in jest that the present disablement of the US presidency over the Iran arms sales affair, and the confusions over US defence policy, obliged Europe to present a united face to Moscow: squabbling in the Community would not help. Moreover, Chancellor Helmut Kohl of West Germany faces general elections next month, so this was not the best time to ask him to be hard on his farming voters. Prime Minister Jacques Chirac of France was preoccupied with the student protests (and late on Friday, the death of a student) in Paris; and the United Kingdom, meanwhile, did not want to bow out of the presidency on a sour note after its efforts these six months to prove itself a hard-working, constructive member of Europe.

The resulting bland unity was not good news for the research programme of the Communities' think-tank and administration, the European Commission, which the Commission has proposed to double. Jacques Delors, Commission president, had hoped the summit could send a strong signal in favour of the programme to the research ministers, who were to meet on Tuesday. During the summit, Delors emerged to say that many delegations had stressed the importance of the Commission's 'framework programme' for research and "approved it totally". Finally, however, the summit communiqué merely "urged" research ministers to agree, but gave no indication on the level of spending on which agreement should be reached, the matter which seemed likely to be the final sticking point on Tuesday between the big three — Britain, France and Germany — which oppose the proposed increases, and the rest, which support them.

President Mitterand, pressed on the research question after the summit — as he had been the first to introduce such matters to summitry — claimed credit that any clause on research at all had been introduced into the communiqué. The British Prime Minister, Margaret Thatcher, also questioned on research, avoided the matter of budgets and stressed the importance of agreement on standards for high technology equipment. She singled out the need for a common standard on digital cellular radio. This relatively small item was also the only technology explicitly mentioned in the communiqué for action, leaving the question of whether the leaders might actually have had in mind the need for common standards in the more ramifying systems of integrated digital services networks.

The European leaders also reached agreement to designate 1989 European Cancer Information Year to develop a "sustained and concerted information campaign . . . on the prevention, early warning and treatment of cancer". No mention was made of possible concerted action on tobacco advertising, but the ministers undertook to "follow up" the recommendations of a committee of European experts which is due shortly to report on cancer prevention to the Commission. A programme on AIDS (acquired immune deficiency syndrome) was also agreed, but this was only to compare notes on actions taken by the different member states in public information and prevention. The possibility of "further cooperation in research" on AIDS was to be considered by health ministers, the leaders agreed. **Robert Walgate**

Independence for UK space centre?

THE British National Space Centre (BNSC) is pressing the government to give it financial independence and to allocate an annual space budget of about £200 million without channelling it through other government bodies. The proposals for the budget and plan are at present being discussed at cabinet level.

The proposal for direct support accompanies the 15-year BNSC plan, and would allow more flexibility in managing the UK space effort. At present, the centre is funded by the Department of Trade and Industry, the Department of Education and Science and the Ministry of Defence.

The BNSC plan would increase spending on space, now £108 million, to about £200 million a year within three or four years. The plan has four principal ingredients and puts heavy emphasis on European collaboration. The first part of the space plan recommends strong British participation in the US Space Station and in a polar platform.

The second plank of the plan will ensure the development of new launchers, independently of the United States. Ariane V would form part of that project together with cheaper forms of space transport, including the European space shuttle HERMES and the British Aerospace/Rolls Royce space-plane HOTOL. The plan also includes an expanded space applications programme — Earth observation, microgravity and the like — as well as more space science research. Studies of solar plasma, geophysics and astronomy would also feature in this programme. Bill Johnstone

Super-collider still fighting for funds

Washington

LAST week the threatened Superconducting Super-Collider (SSC) project cleared one of its major political hurdles by gaining approval from John Herrington, the first Energy Secretary to endorse SSC since the agency began work on the programme five years ago.

It is a change of heart as only a month ago Herrington claimed that his mind was not made up on the advisability of the project. But because of the enormous cost, now estimated to be about $3,000 million, formal approval would have to come from the White House Office of Management and Budget (OMB). According to sources at the Energy Department, Herrington has not yet formally met top OMB officials.

Alvin Trivelpiece, director of the Energy Department's Office of Energy Research, has been trying to get other countries to participate in SSC, but such commitments are difficult if the United States is hesitant. Trivelpiece says Japan may invest $500 million in the project. The Soviet Union is another potential participant. For the time being, potential partners from Western Europe have their hands full with LEP and DESY.

While approval from the Reagan administration is a necessary step for the project to proceed, it is far from sufficient. Many believe that the brake on spending achieved by the current Congress was achieved by adopting financial tactics such as delaying payments and selling assets. Any further reductions will require real programme cuts.

Congressional enthusiasm for the National Institutes of Health shows no sign of diminishing, and the administration is rumoured to be considering a large increase in the budget for the National Science Foundation and a request for more funds for the Strategic Defense Initiative. It is uncertain whether spending on SSC will prove to be any more politically attractive. Joseph Palca

Amazon rainforests

Brazil invites RGS explorers

THIS week the first of 60 scientists depart for Brazil to make the final preparations for the biggest European-led research expedition ever to study an Amazon rainforest. The project team, led by Dr John Hemming, director and secretary of the Royal Geographical Society (RGS), will be based on the island of Maraca close to the watershed between the Amazon and Orinoco basins.

The invitation was made to RGS two years ago by the Brazilian Environmental Secretariat (SEMA), but it has taken since then to organize and raise from a mixture of industrial, private and government sponsors the £125,000 that will fund the basic needs of the year-long project.

Hemming says "we jumped at the invitation since, although Brazil contains almost two thirds of the world's surviving rainforests, it is difficult for foreign scientists to obtain permission to work there."

Half of the teams will come from the United Kingdom with Brazil making up most of the remainder. Maraca, an uninhabited island 60 km by 25 km, is covered in forest, contains a lake and has swamps and seasonally flooded areas. As a result of the expedition, SEMA, whose permanent research station on Maraca is at the disposal of the RGS team, will acquire a complete survey of the island, including an inventory of the animals, insects and plants, to help plan future research.

RGS says its five research programmes have "been welcomed by the Brazilian authorities because of the practical benefit to both farmers and conservationists". They involve an ecological survey, the development of management plans for the reserve, an examination of forest regeneration and the behaviour of soil moisture, a medical-entomological review and a survey of land development.

Caracarahy rapids, Rio Branco

The Royal Botanic Gardens at Kew in London will oversee the ecological survey which will record, classify and collect hundreds of species of trees, plants, vines and ferns with the help of specialists from the Royal Botanic Garden in Edinburgh and the New York Botanical Garden.

There will be an inventory of hundreds of species of fauna, particularly snakes, frogs, bats, monkeys, insects and birds. Studies of forest regeneration could help restore destroyed areas to their original state. According to RGS, thousands of hectares of rainforest are felled every year for farming, settling, cattle ranching and mining. But a wasteland usually results, claims RGS, and much of it is soon abandoned. Farming proves impossible because of poor soils and intense parasite and weed activity. Once damaged, the rainforests do not regenerate and the research will try to find out why.

Another research programme will investigate how the litter (or debris) on the forest floor is derived from the foliage and decomposes, the dynamics of the rainfall as it drops from the roof or canopy of the forest and the behaviour of that moisture with the soil. Research on endemic tropical diseases will focus on onchocerciasis and leishmaniasis.

The land development survey will study the settlements within reach of the research camp. According to RGS, millions of settlers are moving to the frontiers of the Amazonian forests in a desperate search for land. "This occupation and disturbance often occurs too fast for there to be any planning based on scientific information. As a result, many pioneer colonists fail."

Bill Johnstone

Space shuttle

Money-saving move on fuel tanks

Washington

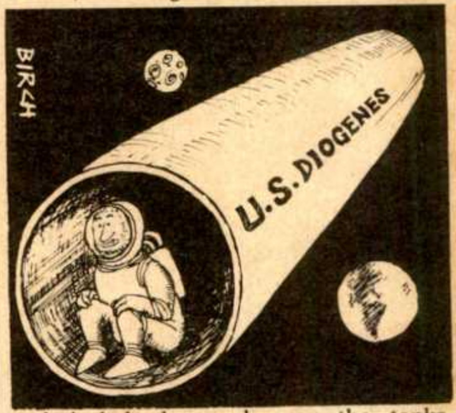
THE idea of making some use in orbit of each space shuttle's huge external fuel tank is not new, but prospects for its realization may have come closer. Hitherto, the fuel tanks have been allowed to burn up in the atmosphere. But now a new company, External Tanks Corporation, wants to make them into warehouses. The company is seeking private capital to finance what it calls the first private "warehouse in space". The project has the backing of the University Corporation for Atmospheric Research (UCAR), a consortium of 57 major research universities. According to Thomas F. Rogers, chairman of External Tanks Corporation and a former member of NASA's space programme advisory council, the first habitable tank could be in orbit by 1993.

The shuttle's external tank carries liquid hydrogen fuel and oxygen for the main engines, and is usually jettisoned just before the shuttle reaches orbit. Its diameter is more than the width of a Boeing 747, and its length is equivalent to a 13-storey building. Twenty-four have so far been discarded, each costing $30 million. To take the tank into orbit would require only an extra 2,000 lb of fuel, according to Randolph Ware, president of the new company.

The company, which is 80 per cent owned by the UCAR Foundation, a non-profit organization established by UCAR, is seeking capital from "companies that have a long-term strategic interest in the exploitation of the space environment", including hotels. The National Aeronautics and Space Administration (NASA) is apparently taking the company's plan seriously. Several studies have recommended that the idea be investigated, and NASA has now formed a special committee to look at it, chaired by general manager Philip Culbertson.

Culbertson says the "very interesting" idea is "feasible", but adds that NASA has not yet looked into all the technical details and that it would in any event leave the business side to External Tanks Corporation. NASA has some plans of its own for using the tanks: it recently awarded a contract to Martin Marietta Corporation to study turning an external tank into a gamma-ray imaging telescope.

Randolph Ware, president of External Tanks, and Rogers, both of whom have technical backgrounds, say the tanks would need minimal modifications on the ground. Once in orbit, the tank would be detached from the shuttle and fitted with a small thruster that would keep it in stable orbit. Subsequently it would be purged, pressurized and fitted as as "the equivalent of high-cost commercial space", according to Ware, who seems to have no doubts that commercial companies will want to rent.



Among the plans being discussed is one for a fully pressurized permanently inhabited laboratory/factory, Labitat, made from two tanks connected together. The operating cost of Labitat is calculated to be so low that accommodation would cost as little as $500 per person per day by the 1990s, roughly twice the cost of first-class hotel accommodation in the United States.

Tim Beardsley

Climatic prediction

Could this be an El Niño?

Washington

SOMETHING is happening in the equatorial Pacific Ocean from the dateline to the west coast of South America. Some believe that the climatic upheaval known as El Niño is in progress, others believe one is ending, some call the events an ENSO episode (El Niño/Southern Oscillation); others claim that all that is happening is that a travelling wave of water is moving through the Pacific from west to east. But there is agreement on one point: something is happening.

The events of interest relate to a warming of the sea-surface temperatures in a band from the dateline eastward, a rise in the water temperature measured from piers along the Peruvian coast, and slackening of easterly sea-surface winds. Whether these events should be termed an El Niño depends on definition. Originally, says George Philander of the National Oceanic and Atmospheric Administration (NOAA), an El Niño referred to warm current flows along the coasts of Ecuador and Peru in January, February and March, and the resulting impact on local weather. But James O'Brien of Florida State University in Tallahassee says no two events have been the same, making the 'classic' definition of little value. The second name, ENSO, more generally refers to events from the mid-Pacific to the South American coast, taking into account the irregular oscillation in pressure between the east and west Pacific.

The National Meteorological Center's Climate Analysis Center (CAC) issued an advisory note on 10 November stating that the warming of sea-surface temperatures in the equatorial Pacific "resembles the warmings associated with ENSO episodes". But the centre was unclear whether El Niño conditions would result. CAC's Don Rodenhouse says the next circular will probably appear next week based on November data, and is likely to confirm that Pacific warming is spreading.

The National Weather Service, CAC's parent agency, is cautious about using the term El Niño. Last March it announced that conditions that could lead to an El Niño were developing. Eugene Rasmusson, now retired from CAC but responsible for the March bulletin, says the warming trend failed to materialize. Many people, mindful of the economic disasters blamed on erratic weather from the 1982–83 episode, mistook the possibility of the event for the event itself, provoking some hyserical behaviour. Now the Weather Service will move very cautiously before making any firm statements about El Niño.

But the government's caution does not extend to the academic community. There is "absolutely no question in my mind that there is a modest El Niño under way", says O'Brien, who is chairman of the ENSO alert group that advises CAC on its Pacific forecasting.

Several members of the group had developed models as early as January predicting the ocean temperatures now observed. But O'Brien says none of them felt that this year's episode would match the 1982–83 episode in scale. From his own model, O'Brien predicts that sea-surface temperatures along the Peruvian coast may soon begin to decline.

Francisco Chavez of Duke University says his data show that decline is taking place based on pier temperatures from Paita, 5° south of the Equator. But Chavez data also show that pier temperatures for the start of November were well above normal.

Tim Barnett of the Scripps Institution of Oceanography has made a statistical model that predicts the behaviour of sea-surface temperature in the equatorial Pacific. Barnett's observations of temperatures in the region show about 1.2 °C above normal. His model predicts that the current warming trend will continue for another month or more.

Mark Cane, Stephen Zebiak and Sean Dalan of Lamont–Doherty Observatory in Palisades, New York, developed a deterministic model based on a coupled ocean–atmosphere system (see *Nature* **321,** 827; 1986). Cane says that if the model is correct, the current warming will have already peaked. His analysis of the latest data from the region supports that view, although the temperature could rise again. Barnett says the exciting thing will be to see how well the models behave in predicting events and to understand why they work as well as they do.

Predicting an El Niño is not necessarily of value in itself. "The term El Niño is useful like the term winter is useful", says NOAA's Philander. The significance of El Niño episodes for global weather is still not well understood. Moreover, Philander is not convinced that the rise in coastal temperature is due to an ENSO episode — a travelling wave, known as a Julian–Madden wave, may be bringing the warm water from the west, rather than a perturbation of the southern oscillation. One measure of the southern oscillation, the pressure difference between the island of Tahiti and the city of Darwin in Australia, has failed to show the difference normally associated with ENSO episodes.

But Rasmusson believes a weak-to-moderate ENSO episode is indeed under way, and the Tahiti-minus-Darwin pressure index is just one of the factors still not clearly understood. **Joseph Palca**

Biotechnology patents

Don't say just what you mean

Rehovot

AN extremely inventive patent covering biologically-engineered "human pink-spider toxinase" was recently presented for examination by Wiezmann Institute's Professor Hermona Soreq and Professor Avigdor Shaefferman of Israel's Biological Research Institute. But the toxinase and the cultured cells yielding a miracle cure for the dread spider bite were all the figments of the inventors' imagination, designed to illustrate the difficulty of patenting intellectual property.

The pink-spider toxinase patent, together with an equally inventive suit claiming patent infringement, was part of an unusual experiment carried out at an international conference on legal aspects of biotechnology held in Jerusalem. The invention was presented for examination by Mr Joel Tzur, the Israel Commissioner of Patents, and it was finally 'approved' after a mock appeal heard by the deputy president of the Tel Aviv District Court, Judge Ellyahu Winograd. Among those attending the Jerusalem meeting were scientists, jurists, patent attorneys, representatives of patent offices and industrialists. The so-called biolegal meeting was the brainchild of Tel Aviv District Court Judge Shoshana Berman. The importance of using precise legal terminology in patent applications became clear at the conference. For example, delegates concluded that it is not appropriate to describe a potentially patentable protein fragment as being of a specific length; instead, it should be described as being *essentially* of a specific length. That extra word makes it far less likely that the application can be successfully challenged. "I learned", Professor Soreq concludes, "that the exactitude demanded of scientists may be counterproductive when they are applying for a patent."

The Israel Patent Office has been issuing patents on genetically engineered microorganisms even before they were patentable in the United States. But Professor Haim Aviv claimed that Israeli authorities have been too liberal in granting broad patent protection to foreign biotechnology companies, which, he charges, are less interested in manufacturing their products than in stifling Israeli innovations.

Aviv, himself deeply involved in the business side of biotechnology, hopes that Israel will soon adopt the US principle of a grace period, so that local researchers, like their US counterparts, will have a full year to patent their findings after publication in a scientific journal.

Nechemia Meyers

Bhopal

Liability for disaster still disputed

New Delhi

THE question of compensation for victims of the Bhopal disaster still remains unresolved on the second anniversary of the leak of methyl isocyanate (MIC). The legal wrangle has taken a new turn with the Union Carbide Corporation (UCC) alleging that it did not own the plant and therefore is not liable for the incident on 2 December 1984.

In a 179-page document filed in the court of Bhopal's district judge, UCC said the Indian claim was based on the "erroneous assumption that the plant belonged to, and was designed operated and maintained" by UCC. It said the plant was designed and constructed by Union Carbide of India Limited (UCIL) in consultation with Indian engineering companies, and that UCC provided only the process design of certain parts of the plant in accordance with the provisions of the design transfer agreement approved by the Indian government. The agreement, according to UCC, specifically stated that UCC "shall not in any way be liable for any loss, damage, personal injury or death resulting from or arising out of the use by UCIL of the design packages". UCC said it was only a shareholder of UCIL which alone "is exclusive owner of the Bhopal plant".

Apart from seeking to transfer responsibility for the plant to UCIL, the US company has denied that the emission of MIC from the plant occurred as a result of defective plant conditions. It said that the incident occurred because large quantities of water were deliberately introduced into the MIC tank by connecting a water hose to the opening from which the tank's local pressure indicator had been removed.

According to UCC, the pressure indicator was in place on 30 November (two days before the accident), but on the morning of the accident it was found to be missing and a water hose was noticed lying on the top of the tank. UCC claims that the plant supervisor's daily notes included a sketch showing that water had entered through a connection to the pressure indicator. It said that somebody deliberately intending to introduce water into the tank could, within a few minutes, close the isolation valve, remove the pressure indicator, connect a water hose from a water service drop only 10 ft away and open the three valves necessary to introduce water.

The case may look straight-forward from down there...
UNION CARBIDE CORP.

UCC said that the facts "establish that a water connection to the tank was found and concealed by the operators and supervisors on the third shift" and that the Central Bureau of Investigation and the scientific advisers, with all the evidence available, "apparently desired not to reach the conclusion that the event was caused by a deliberate act". This "had the effect of aiding the persons involved in maintaining their deception", according to UCC.

The government of India maintain that UCC was liable for the consequences of the incident and is suing the corporation for $3,000 million in damages, saying that 2,347 people died and 30,000–40,000 were seriously injured in the disaster.

The legal arguments are expected to be prolonged, offering no immediate help for the survivors, many of whom, according to the Indian Council of Medical Research, continue to suffer from chest ailments and are without jobs. The Madhya Pradash government has so far spent more than $40 million on health care and rehabilitation; UCC and UCIL have made available about $16 million.

K.S. Jayaraman

Space Station cooperation

Money the root of conflict

Washington

DISAGREEMENT continues over arrangements for European participation in the planned US Space Station. European officials are pressing the United States for assurances on operating procedures and cost-sharing, asking that these be specified in a formal intergovernmental agreement. The United States, however, wants many of the details left to a memorandum of understanding between the National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA). At issue are the financial returns the partners can expect from the collaborative venture and the extent of European autonomy in planning.

Several high-level meetings have been held recently between government and space agency representatives from different countries, including Japan. The negotiations are taking place against a background of uncertainty about the cost of the whole project. Recent NASA estimates suggest that the original figure of $8,000 million will fall far short of requirements unless the station capabilities are reduced. All sides say they want to reach final agreement as soon as possible, frustration is evident. One European official commented that Europe knows from experience that the United States does not take memoranda of understanding to be binding, a reference to the US decision some years ago to scale down its participation in the International Solar Polar Mission.

Officials also concede that serious differences remain over key aspects of management and sharing of costs and space station resources. Europe remains committed to its own $2,000-million pressurized module called Columbus, but is unhappy with the US view that that should limit its share of the total returns to 15 per cent. NASA is required by congressional directive to recoup 80 per cent of space station benefits; one important question, according to Richard Barnes, director of NASA's international affairs division, is whether Columbus should be considered an integral element of the space station system or an adjunct. The same applies to the European plan to develop a man-tended free flier, on which it has insisted despite US misgivings. Questions remain about how shuttle or Ariane launches and provision of payload specialists should appear in the accounts.

Europeans are also unhappy with a US congressional requirement that the United States should be responsible for materials science operations in its modules, leaving biological sciences for the Europeans. Officials say, however, that despite the congressional directive, there is now understanding that Europe will participate somehow in materials science research on the space station and be allowed some of the benefits. It is agreed in principle that commercial companies will have laboratory units that will sometimes be closed to general access in order to keep proprietary information confidential.

Both sides say they want final agreement on both the memorandum of understanding and the intergovernmental agreement by the end of the year or by very early in 1987. The European Council of Ministers is meeting to discuss space next spring, and would like a definite framework established by then; NASA wants to release a final request for proposals for the four main station work packages in January so that detailed development can begin in earnest next summer.

Tim Beardsley

Monoclonal patent

Hybritech versus Abbott

Washington

Hybritech Inc., a subsidiary of Eli Lilly, is suing yet another company for using monoclonal antibody assays in its diagnostic kits. Hybritech has now decided to sue Abbott Laboratories on the grounds that some of Abbott's diagnostic tests may violate its patent for 'sandwich assays'.

The patented Hybritech process uses two monoclonals, each specific to a different site on the same antigen, to make an antibody–antigen–antibody 'sandwich'. This detects small quantities of antigen, allowing early diagnosis of disease.

The interest of the new suit is that Hybritech is asking not only for damages and a permanent injunction against Abbott, but for a preliminary injunction that would prevent Abbott from selling its materials even before the case has come to trial. To persuade a judge, Hybritech will have to prove that continued violation of the patent would cause it "irreparable damage". Patent attorney Jorge Goldstein, of Seidman, Sterne, Kessler and Goldstein, says the chance of winning such an injunction, previously not high, is greater in the present pro-patent legal climate.

The disputed products include Abbott's pregnancy tests and some of its hepatitis and cancer tests, which are sold only to physicians or hospitals and comprise 7 per cent of Abbott's sales of diagnostics. Chuck Weber, a spokesman for Abbott, says that their test for the AIDS (acquired immune deficiency syndrome) virus, which uses polyclonal, not monoclonal, antibodies, is not affected by the suit.

If Hybritech wins this case it could also go after other companies using the sandwich assay technology. Syncor, in Sylmar, California, and Ventrex, of Portland, Maine, both have tests for human chorionic gonadotropin that use Hybritech's sandwich assay process to test for pregnancy. And Centocor, of Malvern, Pennsylvania, has just received US Food and Drug Administration approval for its CA 125 ovarian cancer blood test using the technique.

Hybritech's new suit may have arisen from the latest turn in its earlier suit against Monoclonal Antibodies, Inc. In 1984, Hybritech's suit failed and the patent was ruled invalid due to obviousness, but that decision was overturned, on Hybritech's appeal, last September. This decision returns the case to the federal district court in San Francisco, to determine if Monoclonal Antibodies' products infringe the patent for sandwich assays.

But David Brezner, chief counsel for Monoclonal Antibodies, intends to petition the Supreme Court over the decision upholding Hybritech's patent. There is some possibility that the justices may hear the case, as similar suits will certainly arise over other biotechnology patents in future. Brezner said the Abbott case "will not have a major impact" on his company's litigation with Hybritech.

However the Hybritech cases are settled, their handling by the courts will influence future biotechnology patent cases. Jorge Goldstein predicts a high volume of cases for the coming years. But ultimately, he believes, when several patents have been upheld, fewer companies will be tempted to infringe existing patents.

Carol Ezzell

New hepatitis-B vaccine launched

Brussels

Competition in the potentially lucrative market for vaccines against hepatitis-B virus is hotting up. Following close on the plans of a Korean company to produce a dollar-a-shot plasma-derived vaccine (*Nature* 324, 399, 1986), SmithKline Biologicals last week received official approval (in Belgium) for its new recombinant DNA vaccine. Already extensively tested on volunteers, the SmithKline vaccine will compete directly with a similar recombinant DNA vaccine produced by Merck.

The recombinant vaccines are composed of the coat protein ('surface antigen') of the hepatitis-B virus, made in and purified from yeast. The plasma-derived vaccines are made from particles containing surface antigen that are present in the blood of chronic carriers of the virus. As far as SmithKline has tested it (15,000 doses on 6,000 volunteers), the new vaccine appears to be equally effective and can provide a boost to people immunized with the plasma vaccine. Immunity takes about 4 doses and probably lasts about ten years.

Another antigen that may be important in hepatitis-B virus infection is known as Pre-S; some reports say that antibodies to this antigen prevent the virus attaching to the target cell, others that they predispose a patient to become a carrier. Because of this uncertainty, Pre-S was not included in the current vaccine, but according to Dr Alain Goudeau, who has been conducting clinical trials for the company, they have plans to make and test a vaccine that does contain it.

A course of vaccination with the plasma vaccine at present costs around £90 and SmithKline intends to keep the price levels about the same for the time being; the vaccine has been on sale in Singapore for some months at about 20 per cent less than the plasma-derived vaccine. As the sales increase, it is expected that the price will drop; the company has already invested in a new 1.5 m^3 fermenter to increase volume. SmithKline is confident that it can compete with Merck on price and availability, but the competition with the Korean vaccine on price will probably be the deciding factor.

Rebecca Ward

AIDS research queried

Skulduggery at the lab bench

Washington

"Distrust and dislike" between AIDS (acquired immune deficiency syndrome) researchers is the probable explanation for incidents of tampering at the federal Centers for Disease Control (CDC) AIDS programme laboratory in Atlanta, Georgia, according to a special independent assessment by the Institute of Medicine. The assessment concludes that a "lack of strong scientific leadership" at the laboratory has led to dissatisfaction and low morale among researchers.

The assessment, requested after allegations in the press of data suppression and tampering, says the committee could not rule out minor tampering with equipment. Reports received "suggest that some unusual incidents did in fact occur".

The committee, chaired by Dr Julius Krevans of the University of California at San Francisco, found the AIDS laboratory only "moderately productive" and recommended that CDC make institutional changes to separate basic and applied research. It also says that the director of CDC, Dr James Mason, should "take a substantial and visible role in the implementation of the new arrangements", and called for a review of authorship policy for AIDS research, a "locus of discord". Some individuals said their contributions were not recognized, and others felt that individuals not involved were credited. Some changes have already been made.

Another allegation made in the *Atlanta Journal* and *Atlanta Constitution* that research data were suppressed was not confirmed. The data in question were from experiments investigating the efficacy *in vitro* of Nonoxynol-9, a spermicide, in killing the AIDS virus. Early drafts of a manuscript published in *The Lancet* in December 1985 (Hicks *et al.* **ii**, 1422–23) had been submitted for internal review at CDC in January 1985, but the final version was not submitted until November.

Krevans' committee concludes, however, that although it is understandable that critics wanted to avoid unwarranted extrapolation of the results, "it is not clear why the revision required six months and clearance seven weeks".

Tim Beardsley

Japanese technology

Old computers may prove best

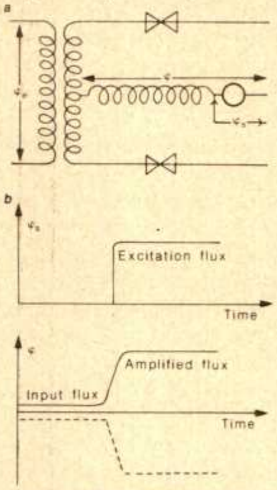
Tokyo

JUST at the time IBM (International Business Machines) drastically scaled down its research and development of Josephson junctions, Professor Eichi Goto of Tokyo University plunged into investigation of the supercooled logic devices, an idea originally developed by him more than 30 years ago. Now, three years later, Goto has been awarded about Y15,000 million ($92 million) by Japan's Research and Development Corporation under the ERATO programme to develop a Josephson junction computer that will process information in the form of minute magnetic fluxes.

The 'quantum magnetic flux logic' project will span five years and will be carried out by three research groups: Hitachi Central Research Laboratory will build the hardware, a team at Mitsui Shipbuilding will develop the software and a group at ULVAC in Chibasaki near Tokyo will be responsible for the magnetic shielding and refrigeration.

Goto achieved worldwide fame in the 1950s with his patenting of the parametron, a resonant electrical circuit that can store binary information by utilizing the two possible phases of oscillation of the resonating current. But although several Japanese companies, including NEC and Hitachi, developed parametron computers, these machines were soon superseded by much faster transistorized computers in the 1960s, and parametron computers can now be found only in science museums.

Nevertheless, Goto believes that a Josephson junction computer using a new type of superconducting parametron could operate at 10 GHz, a hundred times faster than present-day computers, while consuming a thousandth of the power of other Josephson junction devices. The secret lies in using magnetic flux, rather than electrical current, as the conveyor of digital information.



A d.c. flux parametron.

The d.c. flux parametron outputs an amplified magnetic flux the polarity of which is determined by that of a small input signal (see figure). The two possible polarities, positive or negative, can represent the 0s and 1s of binary logic, and by linking parametrons together magnetically, all the basic logic circuits of a central processor can be built. And in fact Hitachi has already linked together three such parametrons to make a clocking device that operates at about 2 GHz.

One major problem, however, remains. The magnetic fluxes on which this computer is supposed to operate are so tiny, on the order of nanogauss, that the computer must be completely shielded from external magnetic fields, such as that of the Earth and underground electric trains.

Undeterred, Goto has already set about trying to build the "perfect magnetic shield". In theory, a superconducting shield encasing the computer could shut out the external magnetic field. But in practice magnetic flux is pinned to the superconducter during cooling, setting up localized internal fields. Goto's plan is to 'sweep' the shield clean of pinned flux with a laser beam, but so far without success.

During preliminary experiments with a superconducting cylinder made of lead, more than 99.9 per cent of the laser's light was reflected by the shiny lead surface and a 'hot spot' for collecting flux failed to develop. Maintaining the cylinder black did not help much either — the paint peeled off on cooling.

If successful, Goto will be able not only to protect the computer from magnetic fields but will also have the ideal detector for magnetic monopoles, which were predicted by the late P.A.M. Dirac.

David Swinbanks

Indian resources

National water policy at last

Bangalore

INDIA has been trying for many years without success to evolve a comprehensive national water policy, but this goal is now a step nearer.

The chief need is for efficient and integrated management of water resources and more effective use of irrigation facilities. Mr B. Shankarananda, Minister for Water Resource Management, says the new policy, which is now getting its final touches, is intended to make available safe drinking water for more than half a million villages and also to link the Himalayan and peninsular rivers so that surplus water can be transferred to areas that need it. Cost-effective techniques for finding water in drought-prone areas, particularly in the north-eastern part of the country, are also being sought.

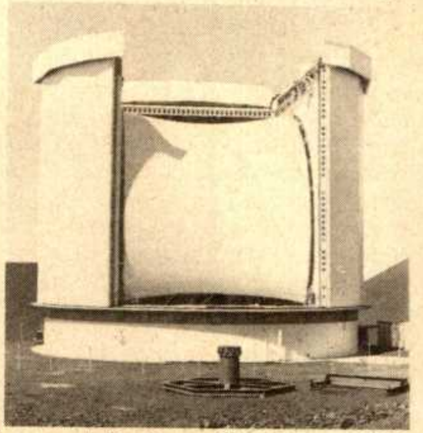
Under the new policy, both central and state governments would share the cost of creating water channels, dams and drinking-water sources. Water would be declared a national asset so as to end interstate disputes.

In the meantime, the Indian Planning Commission has announced that irrigation and flood control projects that do not include plans for 'ancillary aspects' such as soil conservation and protection of the ecosystem will not be permitted.

Water-borne diseases account for the loss of 73 million man-days in India each year. In 1980, it was estimated that out of 570,000 Indian villages, more than a quarter have no source of drinking water within 16 km. Although Rs77 million had been recommended by a working group for rural water supply and sanitation in 1986–90, only Rs34.5 million was allocated.

Radhakrishna Rao

UK–Dutch telescope



LAST week the James Clerk Maxwell Telescope, in the final stages of its initial commissioning tests, received its first radio signals from space. The £20-million instrument, installed at Mauna Kea Observatory on Hawaii, is the largest telescope in the world operating in the submillimetre region of the spectrum.

The telescope, jointly funded by Britain's Science and Engineering Research Council (SERC) and the Dutch Nederlandse Organisatie voor Zuiver-Wetenschappelijk Onderzoek (ZWO) will become fully operational next year.

The first radio signals to reach the telescope last Thursday were those from the Moon, Jupiter and Mars. The telescope has been specifically designed, however, to penetrate the dense regions of interstellar gas, which are opaque to optical radiation. At the infrared and millimetre wavelengths, these regions are transparent.

Bill Johnstone

Italian research reorganization

SIR—Your article on research in Italy (*Nature*, **323**, 751; 1986) depicts the Italian government as working quickly at reorganizing the national research institutions.

As a researcher with 18 years' experience and an employee of the National Research Council (CNR) which, with its 6,189 employees, is the largest of the government research institutions, I would like to express my incredulity at the effects that this difficult reorganization will produce. By the time the reform of CNR is approved by parliament, the most talented and best qualified researchers will have fled to more remunerative positions in universities and industry.

It may be true, as reported in the article, that the average cost per researcher is about 130 million lire. (Please note that public research institutions in Italy employ a mere 14,000 of the 112,884 cited.) But the annual salary of a senior CNR researcher — whose competence, for the information of non-Italian readers, is at the level of a full professor at a university — is no less than 13 million lire, about £6,500.

It is with a sense of shyness mixed with a sprinkling of rage that we watch the CNR president, Professor Luigi Rossi-Bernardi, and the Minister of Research, Senator Luigi Granelli (whose joint efforts to save CNR must be acknowledged), praise the abundant productivity of CNR scientists and the consequent healthy state of CNR. Most of us have spent several years abroad, often working in renowned international research centres, and are all aware of the conditions in which we continue to do research in Italy, and of how they compare with conditions in Britain, the United States and Japan.

Interestingly, the fate of CNR researchers and their hopes of better salaries now seem to be linked to the approval of a government bill (not the reform of CNR) originally devised to exempt the wages of workers at the Mint and of air traffic controllers from the restraints of the anti-inflation finance law, which limits the increase of public employees' wages to 6 per cent a year.

It is well known that Italy suffers from frequent government reshuffles. Experienced observers think it likely that there will be another next March. Whereas I am sure that most of the many "million million" lire projects cited in your article will be funded, I doubt whether the contract for CNR employees (which expired on 31 December, 1984) will be renewed by the end of this year.

MARCELLI G. CACACE
Institute of Protein Biochemistry
and Enzymology, CNR,
Via Tioano, 6,
80072 Arco Felice, Naples, Italy

Maori victory

SIR—Whether or not he was summarizing the views expressed in Crosby's *Ecological Imperialism, The Biological Expansion of Europe, 900–1900*, Roy Porter (*Nature* **323**, 587; 1986) shows extreme ignorance when he suggests that the Maori people of Aotearoa (New Zealand) "were stuck in the time-warp of the Stone Age and could not compete with European arms, ploughs and liquor". He also describes them as having been ruined by syphillis, a suggestion that might be equally appropriately applied to the English monarchy.

The Pakeha (European) myths about the land wars of last century have now been exposed. The British were defeated militarily by the Maori in spite of General Cameron's 18,000 men, the largest standing army at that time in the British Empire. The Maori applied lessons learned during their own musket wars to counter the superior fire power of mortars, rockets, howitzers and the 118-pound Armstrong gun. The gunfighter's *pa* (fortress) consisted of a system of trenches, parapets, rifle pits, covered walkways, fire-proofing and bomb shelters. The success of the Maori military strategy was evident at Gate Pa where Cameron suffered his heaviest defeat and concluded that New Zealand could not be taken by a war of attrition.

A more efficient way of getting Maori land was devised when the Maori Land Court was created and a system of individual ownership was installed to break up tribal lands. The war between Maori and Pakeha over the land and its resources was transferred to the arena of Pakeha-created courts, where it continues today.

The judgement that European domination of the globe is some kind of darwinian success story is premature. Maori resistance to imperialism has been so strong that the Pakeha population of Aotearoa has now been co-opted into the struggle. The parliament of New Zealand is about to pass a law that the governments of both the United Kingdom and the United States interpret as banning their navies from our harbours. We invite Alfred Crosby and Roy Porter to come to New Zealand and expand their view of global biopolitics, but we should warn them that they may catch the anti-nuclear bug which is said to be reaching epidemic proportions in this region.

RANGINUI WALKER
Department of Maori Studies,
PETER WILLS
Department of Physics,
University of Auckland,
Private Bag, Auckland, New Zealand

Pan(da)selectionism

SIR—I have been following with interest the recent correspondence in your pages on evolution, creation science and neo-darwinsim. The letter from T.D. Iles[1], pointing out the regrettable habit of neo-darwinians of making exclusive claims on the interpretation and exposition of evolutionary orthodoxy, was nicely complemented by the Commentary[2] two weeks later on the ancestry of the giant panda. This article, written by a leading (indeed a founding) neo-darwinian, clearly shows the plasticity of neo-darwinian logic.

I do not wish to support either side in the argument over whether the giant panda is more closely related to the raccoon or to the bear (indeed, this is the sort of argument that seems to excite the interest of neo-darwinians in particular: its significance for evolutionary theory in general is less clear), but instead to point out that Mayr's arguments concerning the data on haemoglobin amino-acid sequence differences are circular, self-serving and without experimental proof.

Mayr is concerned on the one hand to prove the irrelevance of the haemoglobin sequence data that show giant pandas to be more closely related to raccoons than to bears, and on the other hand to explain the near identity of the haemoglobin chains of humans and chimpanzees, despite an apparent evolutionary separation of over 10 million years. The explanation offered is, of course, selective pressure on haemoglobin. To quote: "In the case of man and chimpanzee this influence has been conservative, but in the case of the bears it might well have been centrifugal".

What can one say about such *ad hoc* panselectionist arguments that are not supported by biochemical data plausibly relating haemoglobin structure and function to the physiology, habits and behaviour of pandas, raccoons, bears or primates? It is clear that faith in natural selection alone can explain all in these sermons from the pulpit of neo-darwinism. It is surely time to acknowledge that gene and protein sequence data often support theories of molecular evolution in which natural selection of variants is not explanatory[3], or in which common molecular ancestry is disputed[4], and that the use of these data to support either molecular-clock based phylogenetic trees or neo-darwinism selectionist explanations of evolution should be treated with appropriate scepticism.

GREGORY W. WARR
Department of Biochemistry,
Medical University of South Carolina,
171 Ashley Avenue, Charleston,
North Carolina 29425-2211, USA

1. Iles, T.D. *Nature* **323**, 576 (1986).
2. Mayr, E. *Nature* **323**, 769–771 (1986).
3. Dover, G.A. & Flavell, R.B. *Cell* **38**, 622–623 (1984).
4. Schwabe, C. *Trends biochem. Sci.* **11**, 280–283 (1986).

When firsts come only second

Too many contributors to the scientific literature neglect the journalistic principle that the word "first" should prudently be avoided.

Researchers should learn from journalists the dangers of words such as "first" Those who describe some event as the first of its kind are almost certain to be assailed by readers saying that it is nothing of the kind But to play safe by using qualifiers as in "the first recorded demonstration " is merely to blunt the clarity of the message Prudent people therefore stick to verifiable firsts

Here is an illustration of the difficulties that may arise In the issue of *Physical Review Letters* dated 24 November, R P Messmer and P A Schultz from the University of Pennsylvania contribute a fascinating account of their calculations of the electronic structure of C_2F_2, acetylene in which the hydrogen atoms are replaced by fluorine atoms (**57**, 2653, 1086) The interest of the paper is its conclusion that the older valence bond description of atomic bonding may be a more accurate starting point for calculations of the electronic structures of molecules than the molecular orbital framework A corollary is that the electrons participating in the carbon–carbon triple bond are more tightly localized on the two carbon atoms than the textbooks say

Familiar? Even this brief description will prompt many to recall a recent paper by Cooper, Gerratt and Raimondi (*Nature* **323**, 699, 1986) That was also an account of a valence bond calculation of an organic molecule, in that case benzene Again the conclusion was that the valence-bond description can be the starting point for an accurate calculation of the electronic properties Again it emerged that the electrons appear to be more tightly anchored to the carbon atoms than the textbooks say

Both papers are important contributions to a reassessment of the utility of the alternative frameworks for molecular calculations

The nearly simultaneous appearance of very similar pieces of work is not surprising Sometimes, the explanation is that science has developed to the point at which a previously unformulated question cries out for an answer, whereupon several people independently design appropriate experiments

On other occasions, a new technique may be the trigger for several parallel developments, which appears to be the case with the recalculation of the electronic properties of molecules The work of Cooper, Gerratt and Raimondi and of Messmer and Schultz has been made possible by computer programs capable of accommodating the very large numbers of alternative electronic configurations required for accurate valence-bond calculations and the availability of the computer power to implement them

The dilemma of quantum chemistry is cruelly stated by the following string of simplifications Chemical bonding is an essentially quantum phenomenon in which electron spin is crucial, if in the simplest of all natural molecules, H_2, each atom is represented by a proton surrounded by a smeared-out electron but is otherwise dealt with classically, simple electrostatics will suggest that chemical bonding is impossible Only the fact of electron spin allows electron wave-functions to overlap without an intolerable energetic penalty If, for H_2, the spatial wave-function is properly antisymmetrized with respect to an exchange of electrons, the electrostatic energy turns up with a minus sign if the electron spins are in opposite directions, implying chemical bonding in a state of zero total spin

Both the valence-bond and molecular orbital schemes are ways of approximating to the wave functions of molecular electrons In the valence-bond framework, molecular wave functions are built as linear combinations of products of pairs of atomic wave functions, not simply those with one electron attached to each nucleus but also those in which both electrons surround the same proton, leaving the other bare Strictly speaking, excited states should play a part For a complicated molecule, a horrendous number of configurations is usually required

In the molecular orbital view of the world, the problem of H_2 can be dealt with almost exactly by means of the solution of Schrodinger's equation for the hydrogen molecule-ion, two electrons slot with opposite spins into the lowest electronic state, making molecular hydrogen a kind of binuclear helium atom But only modestly more complicated molecules present fearsome problems, the simplest molecular orbital calculations start with approximations to molecular wave functions that are linear combinations of atomic wave functions, and yield useful results only by the empirical fitting of parameters In neither scheme is the set of approximate wave functions an orthogonal set, which leads to further complications Roy McWeeny made the compelling argument, in a *News and Views* article (*Nature* **323**, 666, 1986) that there is no philosophical difference between the two approaches, what matters is merely which is the better starting point for calculation

Much of the interest of what Cooper, Gerratt and Raimondi have done is their use of electronic spin as a way of classifying molecular electronic wave-functions, which is how they conclude that Kekulé's picture of benzene, as a structure in which different patterns of double bonds alternate, may be more realistic than the now-common view of benzene in which there is a static ring of electron density above and below the plane of the six-sided array of nuclei Interestingly, their calculation starts with a set of old-fashioned molecular orbital functions on which they perform what is essentially a valence-bond calculation The outcome is the surprising discovery that electrons are more nearly localized on the six carbon atoms than the conventional picture would have them

Messmer and Schultz take a different tack, starting with the familiar way of dealing with carbon bonds which supposes that the four valence electrons of a carbon atom may occupy any hybrid mixture of the four independent wave-functions of the second electronic energy state The conventional picture of the triple bond in C_2F_2 is that of two links between dumbell-shaped electron orbits, at right angles to each other (called π bonds) and a direct link called a σ bond By cleverly loosening the symmetry restrictions, they are able to show that the state of lowest energy is one in which the bonding electrons at each carbon atom are tetrahedrally arranged, and that the three carbon–carbon bonds are symmetrically arranged about the axis of the molecule Again, the electrons are more localized than expected

These developments are interesting and important The only slightly jarring circumstance is that, admittedly with some qualification, Messmer and Schultz describe their work as a "first", a word whose use is belied by the appearance of Cooper, Gerratt and Raimondi roughly a month earlier) The independence of the two pieces of work, which have only one reference in common, is plain, but the claim to be first makes one seem very mildly foolish Should editors now seek to protect authors from themselves by banning the use of the word "first"?

John Maddox

Solid-state chemistry

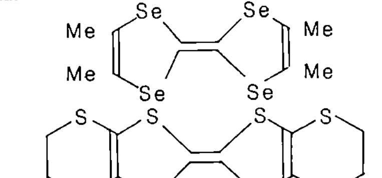
Novel molecular metals

Martin R Bryce

THE scope of materials that display interesting electrical, magnetic and structural properties at very low temperatures has now been extended by Underhill and co-workers, who on page 547 of this issue[1] describe the first two-dimensional organometallic compound that exists in an integral oxidation state The study of the synthesis and solid-state properties of such conducting two-dimensional systems will help to elucidate many fundamental phenomena of solid-state chemistry

Most organic and organometallic solids are electrical insulators, but considerable attention has recently been focused on an ever-increasing number of these solids that exhibit high electrical conductivity and other properties typical of metals This metallic behaviour is caused by extensive intermolecular overlap of systems rich in π-electrons along highly ordered stacks within the crystal The materials are highly anisotropic and are termed quasi-one-dimensional molecular metals[2,3]

Because of instabilities inherent in a one-dimensional electron system, metallic behaviour is almost invariably destroyed at low temperature by a Peierls transition when the materials become semiconducting or insulating (This transition is analogous to the Jahn–Teller symmetry-breaking distortion in molecules) A major goal of research on molecular metals is the stabilization of the metallic state at very low temperatures The report by Underhill and colleagues of a new metal–dithiolene complex with an inherently two-dimensional structure, a feature which suppresses the usual Peierls transition and stabilizes the metallic state down to 1 4 K under a pressure of 12 kbar, is therefore of considerable interest



Top, TMTSF (tetramethyltetraselenafulvalene), bottom, BEDT–TTF (bis(ethylenedithiolo)tetrathiafulvalene)

A significant degree of two-dimensional behaviour has been found previously in molecular metals which are salts of TMTSF and BEDT–TTF (see figure), where X-ray analysis reveals close interstacked selenium–selenium or sulphur–sulphur contacts[4], but two-dimensionality is a new structural feature for highly conducting organometallic salts

The physical basis of the Peierls transition in molecular metals is a direct consequence of the one-dimensional nature of their crystal structures The component molecules crystallize in an orientation such that their electronic states mix to form a continuous band along only one crystal axis The highest energy level within the band that the electrons occupy is termed the Fermi level At low temperatures, long-range order cannot be sustained and lattice distortions lead to the opening of a band gap at the Fermi level When this occurs the conducting chain can be considered to stretch in one region and contract in another, so that the conducting electrons become localized with a filled electron band at a lower energy and an empty band at a higher energy This transition usually occurs at a well-defined temperature (typically between 20 and 100 K) often when the component molecules of the conducting stack dimerize Above this temperature electrons can be thermally excited across the Peierls gap and a uniformally spaced (metallic) stack is energetically preferred

In contrast, for a two- or three-dimensional system, where there is also significant coupling between molecules in a plane perpendicular to the stacking axis, it is less favourable, or even impossible, for the lattice to rearrange and open a band gap The application of pressure to the sample is known to increase the strength of the two-dimensional interactions in TMTSF and BEDT–TTF salts[4]

The discovery by Underhill and colleagues of a complex in an integral oxidation state that nonetheless displays metallic conductivity is also a significant advance When a complex is in a non-integral oxidation state there is only partial charge on each anion This means that energy levels will exist close to the Fermi level that are empty and available to receive the electrons required for conduction along the stack This situation typically arises when there is full charge transfer between donor and acceptor molecules that crystallize in the ratio 1 2 or 2 3, or when there is only partial charge transfer from donor to acceptor that are present in a 1 1 ratio Conversely, for a complex in an integral oxidation state there is a completely filled band and energy is needed to activate electrons into states where they will be available for conduction Such materials are typically semiconductors or insulators In simple terms, movement of electrons along the stack will now necessitate placing two electrons on the same acceptor, with the associated unfavourable coulombic repulsion between these two electrons ☐

1 Parker, I D , Friend, R H , Clemenson, P O & Underhill, A E *Nature* **324**, 547 – 549 (1986)
2 Bryce, M R & Murphy, L C *Nature* **309**, 119 – 126 (1984)
3 Wudl F *Accts Chem Res* **17**, 227 – 232 (1984)
4 Williams J M *Prog inorg Chem* **33**, 183 – 220 (1985)

Martin R Bryce is Lecturer in Organic Chemistry, Durham University, Durham DH1 3LE, UK

Pattern formation

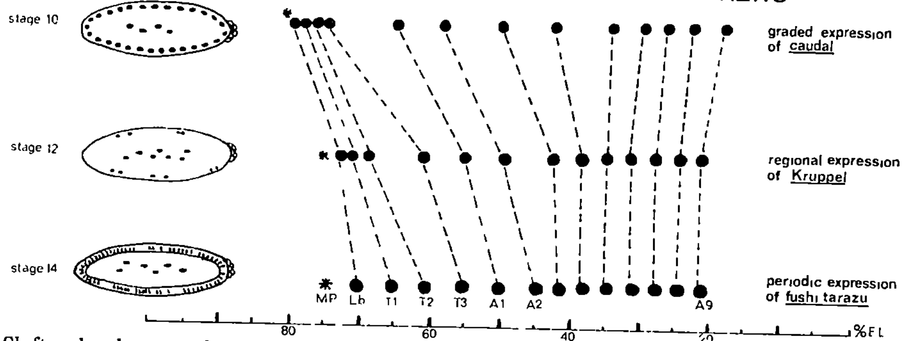
Form and diffusion

Alfonso Martinez-Arias and Philip Ingham

THE question of how cells estimate their location within the body has preoccupied embryologists for the past century, but when posed by Stumpf[1] almost exactly 20 years ago, it seemed to have the beginnings of an answer The concept of gradients, first explicitly propounded by T H Morgan, was entering a new phase of development, and since then it has become almost axiomatic to theories of pattern formation in insects Yet for all that, direct evidence of the existence of gradients has remained tantalizingly elusive The recent discovery[2] of the fruitfly (*Drosophila*) gene *caudal* (*cad*), whose transcripts form a concentration gradient in the early embryo, suggested that the search might at last have borne fruit On page 537 of this issue[3], Macdonald and Struhl report an extensive analysis that argues against a role for *cad* as a primary morphogen In addition, the detailed characterization of primary determinants of polarity in *Drosophila* by Nusslein-Volhard and colleagues[4,5] suggests that a reassessment of the gradient paradigm is timely

Since the beginning of this century, gradients have been invoked as determinants of polarity in a wide variety of organisms[6] Analysis of polarity perturbations in insect segments had in particular provided support for this view In the 1960s both Stumpf[1] and Lawrence[7] realized that the postulated gradients of "unknown diffusible substances"[1] could not only determine polarity, as Morgan had suggested as early as 1905, but also "that a particular level of the gradient might determine the pattern differentiated by cells"[7] at any given position This concept was later embodied in Wolpert's theory[6] of positional information, according to which cells in a morphogenetic field read the local concentration of a morphogen gradient, interpret the information and differentiate accordingly

Although these authors had been con-


stage 10
stage 12
stage 14
graded expression of caudal
regional expression of Kruppel
periodic expression of fushi tarazu
MP
Lb
T1
T2
T3
A1
A2
A9
80
60
40
20
%EL

Shift in developmental capacities of the developing blastoderm (adapted from ref 14) In the stage 10 embryo, when *cad* exhibits a graded distribution, the apparent positions (%EL = egg length) of the segmental primordia (black dots, MP, mouth parts, Lb, labium, T, thoracic, A, abdominal) show a bipartite distribution with clustering towards either pole The shift in these positions between stages 10 and 14 is consistent with diffusion of signals dependent on *osk* and *bcd* Their periodic distribution at stage 14 reflects the expression of *fushi tarazu* and other pair-rule genes

cerned with patterning in secondary fields, the basic concepts of the positional information theory can in principle be applied to the establishment of pattern along the primary body axis during embryogenesis Indeed, Sander's experiments[8] with insect embryos had already led him to propose a role for gradients in generating the insect body plan These experiments excluded the possibility that the cells of the early embryos are specified in response to a fixed set of localized determinants Rather, they encouraged the view that increasing concentrations of morphogens could account for the generation of pattern through the differential activation of genes along the anteroposterior (AP) axis of the embryo[8-11]

At first sight, the *cad* product seems to be a good candidate for such a morphogen Using antibodies raised against the *cad* protein, Macdonald and Struhl[3] and Mlodzik and Gehring[12] find that, like the transcript, the *cad* protein is initially gradually distributed in the early embryo The significance of this distribution is, however, diminished by the results of Macdonald and Struhl's extensive genetic analysis of the locus They have generated mutations that allow them to remove the maternal contribution of the gene from the egg When such eggs are fertilized by normal sperm, the characteristic early protein distribution is barely detectable yet the embryo develops normally

Given the disparity between the levels of *cad* protein with and without the maternal contribution, this result argues that the absolute concentration of the gene product at different points along the AP axis does not have any informational content *per se* Only when the gene is also absent from the zygote is patterning disrupted Thus it seems it is the localized expression of *cad* in the posterior region of the embryo, rather than its initial graded distribution, which is necessary for normal patterning Interestingly, the defects observed are predominantly of the 'pair-rule' type Moreover, they are correlated with alterations in the pattern of expression of the pair-rule gene *fushi tarazu* similar to those caused by the 'knirps-grandchildless' group of mutations[11,13] This suggests that *cad* is involved in the transfer of information between the primary determinants of pattern and zygotic segmentation genes

Some of these primary determinants have recently been identified in a series of elegant analyses by Nusslein-Volhard and colleagues Starting from the premise that the primary determinants of the body plan must be deposited in the egg during oogennesis, her group has isolated maternally acting mutations which disrupt patterning along the AP axis One of the genes so identified, *bicoid* (*bcd*), is required for the normal development of the anterior, whereas another, *oskar* (*osk*), is necessary for normal abdominal development

By carrying out various cytoplasmic transplantations, Nusslein-Volhard and colleagues show that activities dependent on the two genes are initially localized at the anterior and posterior poles of the embryo, respectively In both cases, the mutant phenotype can be repaired by transplantation of wild-type cytoplasm form the appropriate pole In the case of *osk*, the failure to generate the abdominal segments can only be rescued if the posterior polar material is transplanted to the prospective abdominal region of the embryo and not to the posterior pole Normal cytoplasm from the abdominal region does not, however, have rescuing activity This suggests that the material localized in the posterior pole acts as a source which propagates a molecular signal which in some way is transferred to the abdominal region during the early cleavage stages

Similarly, *bcd* activity seems to be localized at high levels in the anterior pole of the egg Transplantation of this anterior polar material not only rescues the *bcd* defect but also can induce anterior structure at any position along the AP axis The concentration of this activity decreases both with distance from the pole and with time, suggesting a dynamic gradient of *bcd* activity in the anterior part of the embryo The status of both gene products as primary morphogens is reinforced by the finding that their transplantation to ectopic sites in the embryo can also induce changes in polarity, its most fundamental property

Despite this compelling evidence for the existence of two sources of diffusible influences emanating from either pole, Nusslein-Volhard's group does not interpret its data within the framework of the clasical gradient paradigm Rather than opposing gradients of *osk* and *bcd* products specifying unique values along the AP axis, both genes are thought to supply more coarse-grained information that would be responsible for the region-specific activation of the zygotic gap genes

In the light of these results, it is interesting to re-examine some of the earlier classical embryological manipulations In 1977 Vogel[14] reported a series of ligature experiments in *Drosophila* analogous to those performed by Sander on other insects On the basis of these, he argued against the generation of the body plan through interacting gradients of morphogens, but instead proposed a regionalization of the embryo consistent with the results discussed above (see figure)

Remarkably, the emerging picture seems not inconsistent with the vision of Driesch[15] who at the turn of this century proposed that "development starts with a few ordered manifoldnesses, but the new manifoldnesses create by interactions, and these are able to provoke new differences and so on With each effect immediately a new cause is provided and the possibility of new specific action"

Although all the new results prompt a re-evaluation of the role and mechanism of action of gradients in primary pattern formation, they do not have any direct bearing on the same aspects of secondary fields originally addressed by Stumpf and by Lawrence Similar experimental approaches will no doubt sharpen our understanding of the latter ☐

1 Stumpf, H *Nature* **313** 430 – 431 (1966)
2 Mlodzik, M , Fjose, A & Gehring, W J *EMBO J* **4**, 2961 – 2060 (1985)
3 Macdonald, P M & Struhl, G *Nature* 324 537 – 545 (1986)
4 Lehman, R & Nüsslein-Volhard C *Cell* **47**, 141 – 152 (1986)
5 Frohnhofer H G & Nusslein-Volhard C *Nature* **324**, 120 – 125 (1986)
6 Wolpert, L in *A History of Embryology* (eds Horder, T J Witkowski J A & Wiley, C C) 347 – 362 (Cambridge Univ Press, 1986)
7 Lawrence, P A *J exp Biol* **44**, 602 – 620 (1966)
8 Sander, K *Adv insect Physiol* **12**, 125 – 238 (1966)
9 Lewis, E B *Nature* **276** 565 – 570 (1978)
10 Nusslein-Volhard C *Wilhelm Roux Arch dev Biol* **183** 247 – 268 (1977)
11 Schupbach T & Wieschaus, E *Wilhelm Roux Arch dev Biol* **195** 302 – 317 (1986)
12 Mlodzik, M & Gehring W J *Cell* (in the press)
13 Carrol, S B *et al* *Nature* **323** 278 – 280 (1986)
14 Vogel O *Wilhelm Roux Arch dev Biol* **182**, 9 – 32 (1977)
15 Stern C *Caryologia* **6**, 355 – 369 (1954)

Alfonso Martinez-Arias is at the MRC Laboratory of Molecular Biology, Cambridge CB2 2QH and Philip Ingham is at the ICRF Developmental Biology Unit, Department of Zoology, Oxford OX1 3PS, UK

Plasma physics

Nuclear fusion comes closer

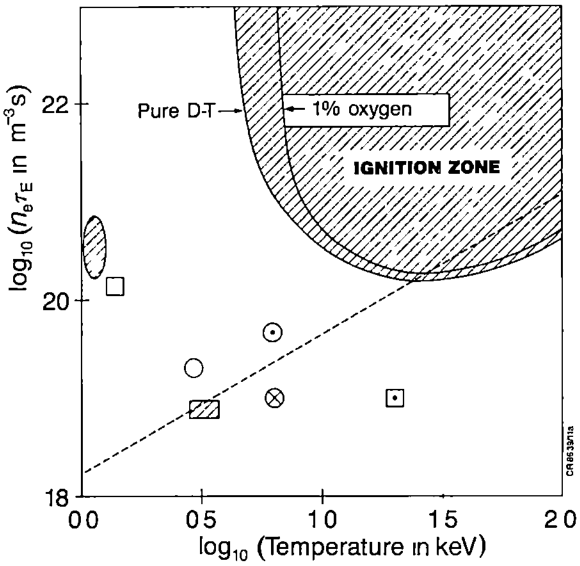
R S Pease

A NEW initiative in international cooperation among the nuclear fusion research community, which may lead to a world engineering test reactor, was announced at a recent meeting* (see below) But the dominant scientific feature of the conference was the experimental results from four large tokamak installations, TFTR (Princeton), D-III-D (San Diego), JET (Euratom, Culham) and JT-60 (JAERI, Japan) All these tokamaks, toroidal electric discharges stabilized by strong magnetic fields, have discharge currents in the range 1 – 5 MA and additional heating in the power range 1 – 20 MW, and are aimed to investigate magnetic confinement of high-temperature hydrogen or deuterium in conditions close to those required for net energy production in a deuterium–tritium (D–T) mixture (see my recent News and Views article, *Nature* **324,** 111, 1986) TFTR and JET, however, are not now planned to operate in D–T before 1990

The results achieved in simple terms of temperature and the critical thermal insulation parameter $n\tau_E$, where n is the ion density and τ_E the energy confinement time, are summarized in the figure, and are compared with the values needed to obtain a self-sustained thermonuclear reaction in a 50 per cent D–T mixture The highest ion temperature (T_i) reported, achieved with the strongest additional heating and at low densities (~ 10^{19} m^{-3}), is about 20 keV, the best thermal insulation, achieved generally with low or no additional heating and at the highest densities, is about 1 5 × 10^{20} m^{-3} s, that is, just at the ignition threshold

The best combination, represented by the triple product $n\tau_E T_i$, is about 2 × 10^{20} m^{-3} s keV obtained both in JET and in TFTR, and is about a factor of five above that reported at the previous conference in 1984, it is still about a factor of ten short of the minimum 3 × 10^{21} m^{-3} s keV required for self-sustained thermonuclear reaction in D–T Up to 10 kW of actual fusion power has been achieved in deuterium alone, which corresponds to more than 1 MW in D–T Typically, such conditions can be sustained for several seconds in JET, the largest tokamak

After allowing for the various differing features of these machines, and for the initial nature of the experiments in D-III-D and JT-60, the major experimental results are all in substantial agreement Moreover, the results from these large tokamaks are supported and supplemented by a wide range of heating and confinement experiments conducted on numerous smaller tokamak installations such as the T-10 (Kurchatov), the DITE and CLEO assemblies (Culham), the HL-1 tokamak in Leshan (China) and high-field tokamaks such as ALCATOR C Generally, all the additional heating methods raise the plasma temperature markedly, have high efficiences and are understood The main phenomenon observed is that the best confinement times are found in high-density operation (Alcator scaling) and in the largest machines But the scaling with temperature and magnetic field is otherwise not very marked



Experimental achievements of ion temperature T_i (generally, the central ion temperature) and density–energy confinement time product ($n\tau_E$) for the four large tokamaks[†] and the NOVA laser compression experiment (E Storm, Livermore) These results are compared with the theoretical values needed for ignition in a D-T plasma (see my review, *Phil Trans R Soc*, in the press) Below the dotted line, the ion and electron temperatures can be significantly different Experimental results of ○, JET (ohmic), ⊙, JET (X-point operation), □, TFTR (ohmic with pellet injection), ⊡, TFTR (best temperature), ⊗, JT-60 (neutral beam heating), hatched rectangle, D-III-D (ohmic), hatched ellipse, NOVA laser compression ([†]R Hawryluk *et al* and S Milora *et al*, Princeton, I Tanga, JET, P -H Rebut, JET, M Yoshikawa, JAERI, and J Luxon, San Diego)

The chief lacuna is that the physical mechanisms underlying the observed energy confinement time remain unidentified Ideally, if the heat transport across the magnetic fields were caused by only ion–ion and ion–electron encounters, then the observed value of τ_E (0 1 – 1 s for large installations) should be about ten times larger, and the corresponding values of $n\tau_E$ should increase as the square of the current In this sense, all four installations have the potential to achieve substantially improved confinement There is growing evidence that this can be done

Pellets

One practical route to better values of $n\tau_E$ is to increase the number density, for example by injecting high-speed (about 1 km s^{-1}) pellets of frozen deuterium This approach is used particularly by the TFTR group and, because τ_E at least does not fall with density (in fact it usually increases), it has met with success which should improve as better pellet injectors are installed Another approach is to adjust the field configuration in the outer regions (X-point operation) This technique, discovered experimentally on the German tokamak ASDEX (Wagner, F *Phys Rev Lett* **49,** 1408, 1982) has met with success in the first reported experiments on JET and on D-III-D The adjustment eliminates the further degradation of confinement time which normally accompanies the additional heating The radial profile of the current density (in the minor cross-section) can be more generally altered by non-inductive current drive up to 1 7 MA was generated by travelling wave radio frequency (2 GHz) fields in JT-60, but the favourable effects on τ_E are restricted so far to low densities Lastly, under a number of conditions, the thermal insulation is observed to show transient improvements localized near the plasma centre, which leads to a strong rise of temperature lasting about a second, followed by a collapse This, again, suggests that long-lasting improved confinement can be obtained

Theoretical studies of the transport problem, stimulated by new, well-diagnosed experimental data, are growing No major instability is, it seems, predicted that could inevitably preclude achieving the required thermal insulation Those that are predicted depend noticeably on the detailed profiles of density, temperature and current, and are governed by the detailed velocity distribution of the ions or electrons and their collision rates None, so far as was reported at the meeting, gives a convincing numerical explanation of the observed heat losses It is quite probable — indeed almost certain — that different mechanisms predominate in different regions of the discharge, so that reducing complex theory to a simple confinement time is no easy task An encouraging development of the experiments is that the impurity levels (~ 1 per cent of O and C) are insufficient to produce a crippling radiation loss

---

*Biennial IAEA nuclear fusion research meeting at Kyoto, Japan, 13–20 November 1986

Taken together, the experiments on the large tokamaks reported substantial improvements in temperature and confinement time arising from the larger dimensions of the new apparatus, from the pellet injection and the powerful additional heating There is a broad empirical basis for expecting further improvement in these parameters to be achieved on these installations, and there are good prospects that the extra experimental diagnostics and the theoretical efforts will lead to the required understanding of the transport process The question of whether even the largest of the installations, JET, can reach partial thermonuclear ignition in D–T fusion is not yet resolved, but tokamak research is clearly flourishing

From the experiments in other magnetic confinement systems reported at the meeting, progress is largely resulting from improvements to already established assemblies For example, a several-fold improvement in confinement time was reported from the HBTX reverse-field pinch experiment at Culham, resulting from careful elimination of defects in the magnetic field system However, the confinement time and temperature in these systems are below those achieved by the strong field stabilization of the tokamak, and are further from the theoretical ideal Both the Japanese (Heliotron E) and German (Wendlestein) groups are operating stellarators using essentially the same apparatus as before Big advances can be expected only when apparatus designed to explore more advanced conditions is brought into operation, for example, new stellarators under construction in Garching, in Canberra and in Oak Ridge, and the reverse-field pinch RFX at Padua

Lasers

Competition for the tokamaks was also reported from inertial confinement experiments, mainly from laser installations in the United States, the Soviet Union, Japan and Europe, but especially from the major new laser installation NOVA at Lawrence Livermore Laboratory, United States and Gekko 12 at Osaka with 70 and 30 kJ output pulses from neodymium glass and phosphate glass, respectively Temperatures of up to 10 keV and densities of up to 50 times liquid were reported Using neutron pulse duration as a measure of τ, the Livermore group reported 2–4 × 10^{20} m^{-3} s as a value $n\tau_E$ at 1 keV, but it was unable to give details of the derivation because its work is classified The direct implosion of D–T contained in thin-walled glass microspheres, where time-profiled radiation pulses were used to mitigate Raleigh–Taylor instabilities, were reported to give the best promise of a net useful energy yield Compression by hohlraum radiation was reported, as also were new target designs using CH_2 foam to stabilize the implosion Both groups reported neutron yields at record values of 10^{13} neutrons per pulse, corresponding to about 10^{-5} burn up of the pellet fuel, and a 2 × 10^{-3} energy yield compared with the absorbed laser energy The laser approach is hampered by the low energy efficiency of lasers, the first engineering tests of the pulsed beam fusion assembly (PBFAII) at Sandia laboratory are therefore important because the charged-particle beams offer the prospect of much higher efficiency The formidable technology problems of 30-MV pulsed transmission lines for nanosecond operation are being overcome, a liquid-lithium ion source based on a dense array of electrostatically induced Taylor cones (J Vandevender, Sandia) is being developed to provide ion emission intensities of 5 kA cm^{-2}

The impression given by these and about 20 papers from smaller installations in France, United States, Britain, China, Canada and Japan is that inertial confinement of high-temperature dense matter is a fast-developing area of physics with many new ideas and results The standard of experiments led by the new US and Japanese facilities is very high

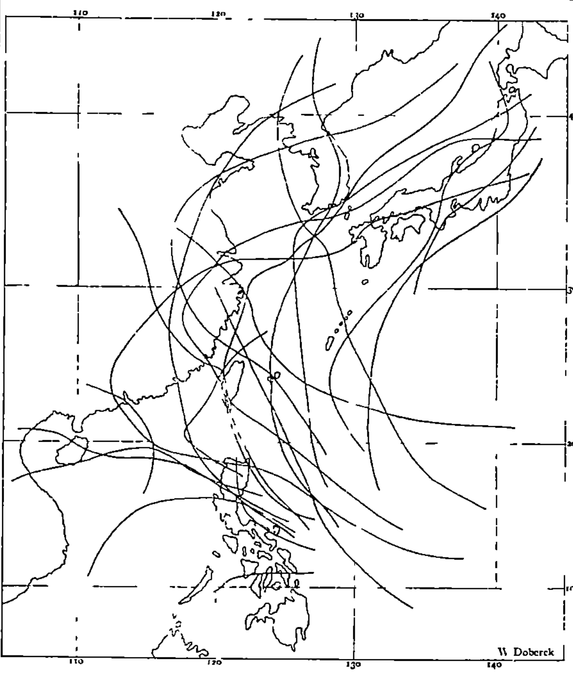
In the sessions devoted to envisaged fusion reactors and reactor technology, one new feature was the first design of a fusion reactor based on the muon catalysed fusion reaction (S Elezier *et al* Israel and Texas), muon catalysis being the dark horse of the fusion race An outstanding feature of the engineering is the successful operation by Kyushu University, Japan of the first superconducting tokamak operating with $NbSn_3$ coils The first experiments with plasma use up to 170 kA and the full fields of 11 tesla (at the coils) There is agreement between experiment and theory on the important issue of shielding the superconducting winding from magnetic transients generated by violent plasma current disruption The engineering problems of the difficult $NbSn_3$ coils, which have delayed completion of the larger Soviet T-15 tokamak, seem to have been overcome in this relatively small scale (major radius 0 8 m) Triam apparatus

The generally bullish outlook on the potential for progress of the magnetic confinement approach to fusion was emphasized by the announcement (J F Clarke, Artsimovitch memorial lecture, 13 November, 1986) of a new proposal for international cooperation Following discussions at the Reykjavik summit meeting, the US government on 30 October transmitted a formal proposal to the Economic Summit nations and to the Soviet Union According to Clarke the proposal is to join together to prepare a conceptual design report (CDR) for an engineering test reactor The proposal indicates that the United States is prepared to commit sufficient resources to accomplish this task within three years If accepted, Clarke believes that this proposal could lead to a major commitment of international resources on preparing for an engineering the most widely accepted element of a common fusion plan ☐

R S Pease is Programme Director for Fusion at the UK Atomic Energy Authority, Culham Laboratory, Abingdon, Oxfordshire OX14 3DB, UK

100 years ago

THE LAW OF STORMS IN THE EASTERN SEAS



From *Nature* 35, 135, 9 December 1886

PRINCIPAL typhoons 1884–1885 Nearly all the typhoons appear to have their origin east or south-east of the high-pressure area that covers the northern Pacific in the summer season, which part of the ocean is characterised by high sea-surface temperature Typhoons are sometimes formed in the China Sea, but then they seldom develop much energy, as they usually move quickly northwards and enter the mainland of China or Formosa Owing to their small dimensions they are easily avoided by such ships as may fall in with them If, however, a typhoon of this kind passes northwards up through the Formosa Channel, it soon becomes as formidable as any of those that originate in the tropical Pacific

Biological diversity

How many species are there?

Robert M. May

ONE of the first things that an extraterrestrial might ask about the planet Earth is how many species are on it. If this extraterrestrial had only our own current knowledge to rely on, the answer would be astonishingly vague: somewhere between 1.5 and 30 million species of plants and animals.

At one end of this spectrum, the number of genera and species of mammals, birds and other large animals are fairly accurately known. As summarized by Diamond[1] in a News and Views article last year, new mammals and birds continue to be discovered (at the rate of about one genus of mammal and three species of birds per year), mainly in the tropics. When we consider the insects and other smaller creatures, however, great uncertainties arise.

Until a few years ago, the uncertainties seemed somewhat less than they do today, and the total number of species was widely believed to lie in the range from 1.5 to 3–5 million. This estimate was obtained roughly as follows. For the species of mammals, birds and other larger animals that are relatively well enumerated, there are roughly two to three times more species in tropical regions than in temperate ones. The total number of species actually named and catalogued is just over one million. Most of these are insects. Indeed, to a good approximation, all species are insects! But most insects that have actually been named and taxonomically classifed are from temperate zones. Thus, if the ratio of numbers of tropical to temperate species is the same for insects as for mammals and birds, we may expect there to be something like two or three yet-unnamed species of tropical insects for every one named temperate species. Hence the overall crude estimate of a total of roughly 3 times the number currently classified, or around 3–5 million.

Estimated numbers of host-specific canopy beetles on *Luehea seemannii*, classified into trophic group (from ref. 3)

Trophic group	Number of species	Estimated fraction host-specific (%)	Estimated number of host-specific species
Herbivores	682	20	140
Predators	296	5	15
Fungivores	69	10	7
Scavengers	96	5	5
Total	1,100+		160

The dramatic upward revision of this estimate results from Erwin's careful studies of the insect fauna in the canopy of tropical trees[2-4]. Erwin used an insecticidal fog (generated by a machine hoisted into the canopy) to 'knock down' the canopy insects; because of their inaccessibility there have been few previous studies of the insect fauna in tropical canopies. Erwin's studies suggest that most tropical arthropod species in fact live in the tree tops, which is not so surprising as this is where there is most sunshine as well as most green leaves, fruits and flowers.

Erwin first collected the arthropods from 19 *Luehea seemannii* trees in Panama over three seasons[2,3]. This tree, a medium-sized evergreen of seasonal tropical forests, has an open canopy and large and well-spaced leaves. The trees chosen for sampling had few epiphytes or lianas, and yielded more than 1,100 species of canopy-dwelling beetles (including weevils), distributed among the categories of herbivore, predator, fungivore and scavenger as shown in the table. To use this information as a basis for estimating the total number of insect species in the tropics, one needs to know what fraction of the fauna are specific to the particular three species or genus under study. Unfortunately, as Erwin emphasizes[3], "No data are available with which to judge the proportion of host-specific arthropods [on trees] per trophic group anywhere, let along the tropics. So, conservatively, I allowed 20 per cent of the *Luehea* herbivorous beetles to be host-specific (that is, they must use this tree species in some way for successful reproduction), 5 per cent of the predators (that is, tied to one or more of the host-specific herbivores), 10 per cent of the fungivores (that is, tied to fungus associated only with this tree), and 5 per cent of the scavengers (that is, are associated in some way with only the tree or with the other three trophic groups)." The results summarized in the table (which is slightly modified from Erwin's work in that I thought it inappropriate to go beyond two-figure accuracy in the last column), suggest about 160 host-specific canopy beetles per genus-group of trees.

Several other assumptions are needed to extrapolate from the estimate of 160 host-specific species of canopy beetles per tree to 30 million species in total[3,4]. Slightly simplified, the argument runs as follows. First, Erwin notes that beetles represent roughly 40 per cent of all known arthropod species, leading to an estimate of 400 canopy arthropod species per three species. Next, he suggests the canopy fauna is at least twice as rich as the forest floor fauna, and is composed mainly of different species; this increases the estimate to around 600 arthropod species that are specific to each genus-group of tropical tree. Citing an estimate of 50,000 species of tropical trees[5], Erwin thus arrives at the possibility that there are 30 million tropical arthropods in total.

It is easy to cavil at each step of this chain or argument. For instance, the fact that 40 per cent of taxonomically classified arthropod species are beetles is of doubtful relevance if in truth essentially all arthropods are unclassified tropical canopy dwellers; what we need to know is what fraction of this canopy fauna consists of beetles. Furthermore, the overall estimate depends almost linearly on the necessarily arbitrary assumption that 20 per cent of the herbivorous beetles are specific to a genus-group of tree; changing this number to 10 per cent would halve the global estimate to 15 million species. Nor is it clear to me that each of the 50,000 species of tropical trees constitutes a 'genus-group' in the sense used by Erwin. These reservations do not, however, detract from my admiration of his work, which has advanced us to the point where we can begin systematic investigation of each of the links in his chain of argument; in comparison, the earlier estimate of 3–5 million is pure hand-waving.

There is a different argument, altogether independent of any of those I have discussed so far, which suggests the total number of species exceeds 10 million. A plot of the number of species as a function of the physical size of the constituent individuals (on a log–log plot) gives an almost linear regression line, with a slope around −2, for species with characteristic linear dimensions above 1 cm or so[6]. That is, for the relatively well-catalogued larger species a tenfold reduction in linear dimensions results in roughly a hundredfold more species; there are roughly 100 as many species of characteristic size 10 cm than of size 1 m. But as we move below 1 cm or so, the number of classified species falls further and further below the regression line that fits larger size categories[6]. Were this regression extrapolated down to around 0.5 mm, we would increase the species total from the known 1 million or so to more than 10 million. As we lack a fundamental understanding of the size-species relation itself, there is no reason to expect such a simple extrapolation to estimate accurately the number of unclassified smaller species, but it is interesting that the number is not inconsistent with

Erwin's more biologically based estimate.

Moving on to smaller organisms, we have to consider protozoa, bacteria and viruses. Some would argue that the species concept breaks down when applied to viruses and bacteria, with their recombining plasmids and the like. Others observe that the small number of species of 'microparasites' currently classified are disproportionately pathogens of *Homo sapiens*, and that a highly conservative estimate would be one such microparasite species particular to each species of invertebrate or vertebrate animal or plant. This would double the total species count at a stroke. Whether anyone cares for the preservation of these microparasite species is another question. A touching note in the Smithsonian Museum of Natural History records the passing in 1914 of Martha, the last passenger pigeon; no-one wept for the passing of the last wild smallpox virus. Many popular ecology books explain how predators help the prey species by culling diseased individuals, but I have yet to see the suggestion that microparasites are beneficial in supplying weakened prey to sustain predator populations.

In between the furries and featheries that excite compassion and the viral, bacterial and protozoan pathogens that we try to eliminate lie the insects, to which both taxonomists and conservationists have been relatively indifferent. Erwin's work is important in focusing concern on this huge class of animals that has been neglected for too long. □

1. Diamond, J.M. *Nature* **315**, 538 (1985).
2. Erwin, T.L. & Scott, J.C. *Coleopt. Bull.* **34**, 305 (1980).
3. Erwin, T.L. *Coleopt. Bull.* **36**, 74 (1982).
4. Erwin, T.L. *Bull. Ent. Soc. Am.* **2**, 14 (1983).
5. May, R.M. in *Diversity of Insect Faunas* (eds Mound, L.A. & Waloff, N.) 188–204 (Blackwell, Oxford, 1978).

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544, USA.

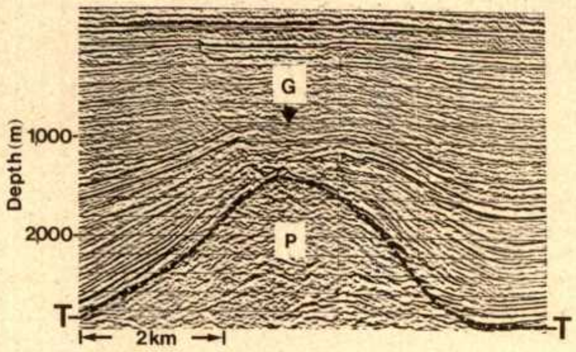
Halite rheology

Salt with a pinch of water

Malcolm K. Jenyon

THE connection between the rheological behaviour of the mineral halite (NaCl) in the presence of very small quantities of free water and radioactive waste disposal may at first seem obscure. There is a strong link, however, some aspects of which are described by Urai *et al.*[1] on page 554 of this issue. Adding to previous work[2,3], these authors go some way towards elucidating the marked effects on halite rheology of the very small quantities of brine inherent in a polycrystalline salt mass. The investigations have important implications for the determination of long-term stability in any rock salt body that is being considered as a repository of radioactive waste or for material storage. Their findings are also of importance in the wider field of salt tectonics.

Natural deposits of rock salt, the main constituent of which is halite, usually occur as members of evaporite cycles in stratified layers or as the thickened 'pillows' and piercement structure (diapirs) resulting from plastic deformation of the source salt layer. These salt masses are attractive as sites for material storage or waste disposal for two main reasons: first, rock salt at depth is impermeable to the types of material involved; and second, by solution mining it is relatively easy to create storage caverns in the salt of sufficient volume to be economically viable. Statistics of usage are not easy to come by (particularly in the case of radioactive waste), but as an example of use for hydrocarbon storage, Dreyer[4] mentions that about 140 caverns (most being in salt bodies) with a total volume of 22 million cubic metres were either completed or under construction in the Federal Republic of Germany in 1982.



Seismic section through a salt pillow (P) produced by plastic deformation of the salt source layer, the top of which is indicated (T). The overburden above T consists of Mesozoic and Tertiary clastic sediments. At G a collapse-graben is seen, resulting from subsidence of the overlying sediments into the incipient void following dissolution and removal of salt by under-saturated water circulation in overburden fractures above the pillow crest, and in the crestal zone of the salt feature. Thickened salt bodies of this type, and the diapiric (piercement) structures into which they frequently evolve, may be considered as sites for cavern construction and the storage/disposal of various materials including radioactive waste. Scales are approximate; vertical scale exaggerated approximately 2.5:1.

In the case of radioactive waste disposal, it is clear that determination of the long-term stability of a salt cavern is a priority for reasons of safety, needing a wide range of investigations to study the tectonic framework of the area (including the fault systems and local seismicity) and to measure the stability of the bedrock beneath the salt; that of the groundwater system; the permeabilities of salt and 'country rock'; the geothermal gradient; the internal integrity of the salt body itself; and the stress patterns and movement history of the salt body (see figure).

The question of the movement history of the salt body is important in assessing the long-term stability of any cavern that is constructed; however, the rheological behaviour of halite is still imperfectly understood. It is known that over a range of temperatures far below its melting point of about 800 °C, halite deforms plastically by intragranular dislocation and gliding, and probably by diffusive mass transfer (pressure solution) through thin, intergranular aqueous films. Increasing the temperature reduces the yield strength of halite, as does the application of differential pressure, although increasing the confining pressure seems to have a much smaller effect. The presence of free water also reduces the yield strength of halite; however, as a polycrystalline halite mass buried to depths greater than about 300 metres is thought to be impermeable (see, for example, ref. 5), the effect of free water on plastic deformation in buried salt masses has been little studied.

The work of Urai *et al.*[1] shows that water need not enter a salt body from outside for the weakening effect of free water to be a factor. The authors demonstrate that even the minute quantities of inherent brine in a salt body (that contained in fluid inclusions in the halite crystals, and particularly in intergranular films) may have profound effects on the rheology of salt. Of particular importance are indications that, whereas at higher strain rates the deformation mode is that of dislocation/gliding, at very low strain rates there is a transition to pressure solution, by grain boundary diffusion with dynamic recrystallization along the boundaries. Also, lattice diffusion can occur, with water molecules passing through the NaCl lattice, and this, together with the brine inclusions

1. Urai, J.L., Spiers, C.J., Zwart, H.J. & Lister, G.S. *Nature* **324**, 554 (1986).
2. Spiers, C.J., Urai, J.L., Lister, G.S., Boland, J.N. & Zwart, H.J. *European Atomic Energy Commnity Contract WAS-153-80-7-N(N)* (Inst. Earth Sci., Univ. Utrecht, in the press).
3. Spiers, C.J., Urai, J.L. & Lister, G.S. in *Proc. 2nd Conf. Mechanical Behaviour of Salt 1984* Hannover (in the press).
4. Dreyer, W. *Underground Storage of Oil and Gas in Salt Deosits and Other Non-Hard Rocks* (Halsted, New York, 1982).
5. Baar, C.A. *Applied Salt-Rock Mechanics 1* (Elsevier, Amerstam, 1977).
6. Odé, H. *Am. As. Petrol. Geol. Mem.* **8**, 70 (1968).

available within halite crystals, may affect the dynamics of the dislocation mode of deformation.

The findings go a long way towards explaining apparent inconsistencies (based on current ideas) in salt deformation at different depths and in different areas of some salt basins. They may also explain unusual phenomena such as the marked increase in flow rate during the wet season of the rock salt under normal temperature and pressure at the surface in the emergent Hormuz salt diapirs of Iran, and also the Joffé effect, in which a halite crystal exhibits a marked increase in ductility when deformed under fresh water[6].

Plainly, the results obtained so far demand that attention is paid to the amount and distribution of inherent brine in the rock salt members of evaporite sequences in any basin, to make an effective analysis of the causes and results of plastic deformation. The message that low-stress conditions acting over long periods can produce significant plastic deformation in salt rock will have a considerable impact on our ideas about salt deformation in general. □

Malcolm K. Jenyon is Manager, Interpretation, at Seismograph Service (England) Limited, Holwood, Keston, Kent BR2 6HD, UK.

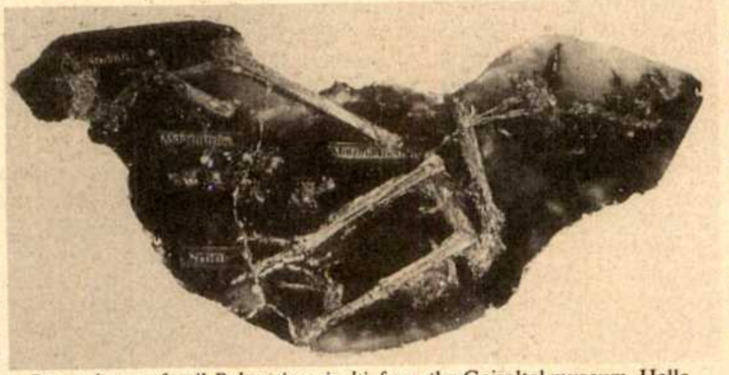
Phylogeny

The German 'ostrich' and the molecular clock

Michael E. Howgate

ONE of the major problems with molecular clocks is that they need to be located in an absolute time frame determined by the fossil record. Undoubtedly the best such clock available at the moment is the DNA–DNA hybridization clock of Sibley and Ahlquist[1], but Peter Houde on page 563 of this issue[2] argues that the DNA–DNA clock needs to be reset.

This molecular clock takes as one of its reference points the divergence of ostrich DNA from rhea DNA. This divergence has been timed from the break-up of the super-continent Gondwanaland which isolated the supposedly already ground-dwelling ratites (ostrich, rhea, emu, cassowaries and kiwi) on the different parts of the fragmenting supercontinent during the late Mesozoic. The palaeontological basis for such a reference point has always been weak as fossils ascribable to the modern genera of the ratites have not been found in strata lower than the Miocene. So it is purely speculative that the ratites were grounded before the break-up of the Gondwanaland supercontinent.


Boney issues: fossil *Palaeotis weigelti*, from the Geiseltal museum, Halle.

The ratites are a rather odd assemblage of ground-dwelling birds confined to the southern continents and with a basic overall similarity due to their adoption of a flightless condition. What separates them from most other birds, even some which have become flightless, is the relatively primitive, reptilian (palaeognathous) structure of the palate and the presence of rhamphothecal grooves in the mandible. The ratite paletal structure, which is shared with the tinamous (themselves poor fliers) of South and Central America, is intermediate between the palate of archosaurian reptiles and that seen in all other birds — the neognathous condition. This 'primitive' avian palate involves a characteristic arrangement of skull bones, with comparatively large vomers and palatines, the latter being firmly united to the pterygoids producing a less flexible skull than that observed in neognathous birds.

The current consensus of palaeontological and biochemical data indicates that ratites and tinamous are a monophyletic assemblage, although whether they arose from a neognathous ancestor or are primitively palaeognathous is still a matter of contention. Olson[3] outlines five possibilities to account for the palaeognathous condition of these birds: (1) that they are monophyletically derived from neognathous birds; (2) that such a derivation is due to neotony; (3) that their derivation is completely separate from the neognathous birds, with the different ratites evolving from a 'proto-tinamou'; (4) some living palaeognathes are derived from a primitively palaeognathous stock, whereas others are derived from one or more neognathous stocks; and (5) all living palaeognathes are derived from different neognathous stocks.

Houde now suggests that an assemblage of palaeognathous birds with well-developed wings and pectoral girdles — the *Linothis*-cohort — are the direct ancestors of all living ratites and evolved distinctive features of the ratite palate and bone histology during the Palaeocene and early Eocene in Europe and North America. This confirms the third of Olson's suggestions except that according to Houde the ratites diversified from a post-tinamous rather than a pre-tinamous condition. Houde further hypothesizes that a bird that retained the power of flight at this time must have flown to New Zealand to give rise to the extant kiwis, which retain some of the primitive characteristics of *Linorthis* that have been lost by other more derived ratites.

Meanwhile in Europe the *Linorthis*-cohort gave rise to an ostrich-grade palaeognath *Palaeotis weigelti* during the middle Eocene (see figure). Thus, ostriches appear to have evolved in Europe and subsequently emigrated overland to Africa at any time from the late Eocene onwards.

If the picture envisaged by Houde is correct and flying ratite ancestors inhabited North America and Europe during the Palaeocene and Lower Eocene and gave rise to an ostrich-grade descendant in the middle Eocene in Europe, long after the break-up of Gondwanaland, then the datum on which DNA–DNA hybridization data is tied in to the fossil record is out by tens of millions of years.

What may be true for ostriches may not apply to the other ratites, and although Houde does not mention it he must hypothesize an almost world-wide distribution for his *Linorthis*-cohort to explain the distribution of the ratites, and he does suggest they must have flown to New Zealand. The lack of fossil evidence for *Linorthis et al.* in the southern continents may only be Northern Hemisphere collecting bias and misidentification of the sort which plagues avian palaeontology. How else is one to explain the presence of the presumed flightless palaeognath *Diagenornis fragilis* from the Palaeocene of Brazil? The problem of the fossil record, however , does not conflict with his critique of the DNA–DNA clock; it is clear the calibration will require further scrutiny. □

1. Sibley, C.G. & Alquist, J.E. in *Evolution Today, Proc. Second Int. Congr. Syst. evol. Biol.* (eds Scudder, G.G.E. & Reveal, J.L.) 301 – 335 (Carnegie-Mellon Univ.).
2. Houde, P. *Nature* **324**, 563–565 (1986).
3. Olson, S.L. in *Avian Biology* vol. 8, 80 – 238 (Academic, New York, 1985).

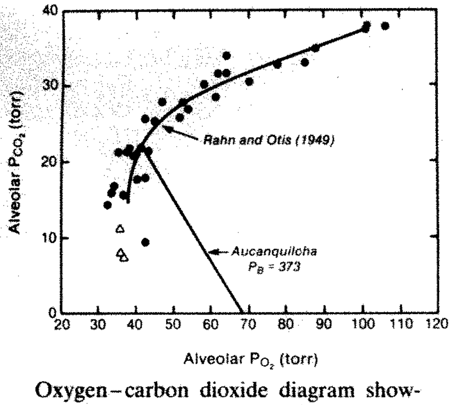
Michael E. Howgate is in the Department of Zoology, University College London, London WC1E 6BT.

Highest inhabitants in the world

SIR—In 1935, members of the International High Altitude Expedition went to the Aucanquilcha sulphur mine in north Chile (21°12'S, 68°28'W) to study their own physiological adaptations to extreme altitude, and also to make some measurements on the local residents[1]. They described a group of miners who worked at the mine, which was said to be at an altitude of 5,800 m, but lived at an altitude of 4,340 m, preferring to climb 500 m on foot every day, which took about an hour and a half, rather than live at the extreme altitude of the mine.

The fact that the miners did not live at the mine itself has often been cited as evidence that people cannot live that high. Furthermore when eight physiologists on the Himalayan Scientific and Mountaineering Expedition of 1960–1961, under the scientific leadership of L.G.C.E. Pugh, spent up to 100 days in a prefabricated hut at an altitude of 5,800 m where the barometric pressure averaged about 380 torr, they seemed to confirm the view. In spite of reasonably good living conditions, ample food, and a healthy exercise programme, they steadily lost weight at the rate of 0.5–1.5 kg per week[2].

It is remarkable that although the Aucanquilcha miners have often been cited as the highest inhabitants of the world[3,4], and therefore are of considerable scientific interest, there is no report of them being visited by physiologists since 1935. Last year, however, I had the opportunity to visit the mine with the Chilean physiologist, Raimundo Santolaya, and was astonished to find that four caretakers were living at the mine itself in a galvanized iron hut. One man had been living there for two years when I met him, but most of the caretakers live there for shorter periods and they all descend by truck to Amincha (4,200 m) on Sundays to play football. Conversations with the manager of the mine made it clear that only a small subset of miners are able to tolerate the altitude. Nevertheless it is clear that some people can live at this altitude for indefinite periods of time. The miners are mostly Bolivian natives who normally reside between 2,000 and 4,000 m.



Oxygen–carbon dioxide diagram showing alveolar gas composition of both permanent residents and lowlanders acclimatized to high altitude. Most of the data were collated by Rahn and Otis[6]. The three triangles bottom left were obtained at extreme altitudes on Mount Everest[7]. Diagonal line shows respiratory exchange ratio of 0.85 for barometric pressure of 372 torr and predicts that the inhabitants at 5,950 m will have alveolar P_{O_2} and P_{CO_2} values of approximately 42 and 22 torr, respectively.

Although the 1935 expedition put the mine at 5,800 m, a 1972 large-scale contour map of the area (Institute of Military Geography, Santiago) puts it at 5,950 m. We measured the barometric pressure as 373 torr and a reading of 371 torr was made by R.B. Schoene in March 1986. The pressure–altitude relationship is latitude-dependent and the latitude of Aucanquilcha at 21° south is similar to that ot Mount Everest (28° north) where the relationship has been accurately measured[5]. This confirms an altitude of approximately 5,950 m.

Few physiological data are yet available on these men. The barometric pressure of 371–373 torr means that the P_{O_2} of moist inspired gas is only 68 torr. If their alveolar P_{O_2} and P_{CO_2} conform to the pattern previously described for acclimatized lowlanders and high altitude residents[6,7] they would have alveolar P_{O_2} and P_{CO_2} values of approximately 42 and 22 torr respectively (see figure). Arterial blood samples were taken from three of the men in March of this year when they were visiting Ollague (3,700 m P_B 490 torr) and these gave a mean haematocrit of 60.89 per cent which is similar to that of acclimatized lowlanders at a similar altitude[3].

Few people live permanently above 5,000 m even in the Peruvian Andes, or the Himalayan ranges of Nepal or Tibet. The caretakers of the Aucanquilcha mine apparently live at a far greater altitude than anyone else in the world.

I am indebted to R. Santolaya for making the arrangements to visit the mine, to R.M. Winslow for measuring the hematocrits and to R.B. Schoene for a measurement of barometric pressure.

JOHN B. WEST

Department of Medicine,
University of California,
San Diego, La Jolla,
California 92093, USA

1. Keys, A. *Sci. Montly* **43**, 280-312 (1936).
2. Pugh, L.G.C.E. *Brit. med. J.* **2**, 621-627 (1962).
3. Ward, M. In *Mountain Medicine* (van Norstrand Reinhold, New York, 1975).
4. Heath, F. & Williams, D.R. in *Man at High Altitude* (Churchill Livingstone, London, 1977).
5. West, J.B., Lahiri, S., Maret, K.H., Peters, R.M. & Pizzo, C.J. *J. appl. Physiol.* **54**, 1188-1194 (1983).
6. Rahn, H. & Otis, A.B. *Am. J. Physiol.* **157**, 445-462 (1949).
7. West, J.B. *et al. J. appl. Physiol.* **55**, 678-687 (1983).

Functions of leaf fall

SIR—Ford has suggested that leaf fall is the mechanism of excretion in vascular plants[1]. But to view the senescing leaf simply as an 'excretophore' may be to limit our vision. Some species may avoid accumulating high concentrations of heavy metals by removing them in senescent leaves[2]. The dead leaves of other species release allelopathic compounds which play an active part in interspecific competition[3]. And the maintenance — indeed the creation — of the plant's habitat, the A and B soil horizons, depends on leaf fall. The degree to which nutrient elements are withdrawn from leaves before they fall varies between ecological communities and correlates with the availability of nutrients in the soil[4]. Thus elimination of waste products of metabolism may not be leaf fall's only adaptive function.

A contrary view[5] is that higher plants have been able to derive a variety of selective advantages from their dead and dying parts, thanks to their 'modular' form of construction[6]. This allows parts of the organism to be shed without prejudicing the survival of the whole. The result is that as in life so in death and fall, leaves have been subject to many selection pressures and many functions are now achieved by the process.

The practical importance of all this is that it implies that leaf senescence is not simply an avoidable evil. The use of 'anti-senescence' genes or chemicals to increase leaf area duration and therefore yield in agricultural crops, may prove to be not quite as straightforward as has sometimes[7] been assumed.

RICHARD C. HARDWICK

Institute of Horticultural Research,
Wellesbourne, Warwick CV35 9EF, UK

1. Ford, B. J. *Nature* **323**, 763 (1986).
2. Waughman, G. J. & Bellamy, D. J. *Ann. bot.* **47**, 141-145 (1981).
3. Sydes, C. & Grime, J. P. *J. Ecol.* **69**, 237-248 (1981).
4. Grubb, P. J. *A. Rev. Ecol. Syst.* **8**, 83-107 (1977).
5. Hardwick, R. C. *British Plant Growth Regulator Group Monogr.* **9**, 61-74 (1983).
6. Hardwick, R. C. *Phil. Trans. R. Soc.* **B313**, 161-173 (1986).
7. Abu-Shakra, S. S., Phillips, D. A. & Huffaker, R. C. *Science* **199**, 973-975 (1978).

Undirectional spin

SIR—When I wrote my paper on undirectional spin[1], which was recently referred to in *Nature*[2], I unhappily had not seen a paper[3] on this topic which preceded mine by a few years. This highly illuminating work, which overlaps mine hardly at all, is particularly to be recommended to your readers for its excellent simulations.

HERMANN BONDI

Churchill College,
Cambridge CB3 0DS, UK

1. Bondi, H. *Proc. R. Soc.* **A405**, 265 – 274 (1986).
2. Walgate, R. *Nature* **323**, 204 (1986).
3. Kane, T.R. & Levinson, D.A. *Int. J. non-lin. Mech.* **17**, 175 – 186 (1982).

Detecting Earth-like planets

SIR—Angel *et al.*[1], while advocating the use of a 16 m diameter telescope optimized for infrared observations to search for Earth-like planets outside the Solar System, criticize my suggestion[2] of using an interferometric system, operating at optical wavelengths, for the same purpose. Their criticism supposes that the sidelobe levels would be objectionably high because of the optical imperfections in the interferometry elements but here we point out that there is a fundamental error in their analysis.

Sidelobe suppression is an important issue for both concepts. The interferometer that I discussed would consist of a number of telescopes of diameter D, deployed in an array with spacings larger than D, but not so large that a planetary disk would be resolved. The elements would be pointed away from, but still close by, the star in searching for planets, but photons from the star would cause a noise background because of the finite sidelobe response. The star could be used as a phase reference to stabilize the interferometer, thus allowing coherent integration, and the diffraction sidelobes would be minimized by apodizing the aperture illumination of the individual elements (which were 1.5 m telescopes in the example given). Still, there would be noise sidelobes introduced by surface irregularities in the optics and it is control of these sidelobes that was being questioned by Angel *et al.*

To summarize the causes of noise sidelobes, the usual definition of gain will be used; relative gain (sometimes called directivity) is the gain normalized to unity on the optical axis (this is the inverse of the definition given by Angel *et al.*). The gain contribution from the noise sidelobes at an angle (θ) off the optical axis will be denoted by $G_n(\theta)$. Intensity in the image plane will be the Fourier transform of the auto-correlation of the amplitude error in the aperture-plane phase front. The precise form of the sidelobes will depend upon the error distribution, but as a representative case let the errors have a random distribution, with an auto-correlation of the form

$$<y^2(\tau)> = <\delta^2> 1 - e^{-(2\tau/l)} \qquad (1)$$

for separation τ: ι is a correlation length and is the phase error. If the r.m.s. surface error is ε, then $<\delta^2> = (4\pi\varepsilon/\lambda)^2$. It then follows from the work of Ruze[3] that, for a uniformly illuminated aperture of diameter D with $\delta << \lambda$, the relative gain of the noise sidelobes will have the distribution

$$G_n(\theta)> = 8\pi^2(\iota/D)^2(\varepsilon/\lambda)^2 e^{-(\frac{1}{2}\pi l\theta/\lambda)^2} \qquad (2)$$

Physically, the gain distribution will be a speckle pattern extending over an angle $\approx (\lambda/l)$ with individual speckles of size $\approx (\lambda/D)$.

In their remark that "The speckles will look like a cluster of barely resolved stars of 21st magnitude, in which is hidden a 28th magnitude planet" Angel *et al.* misunderstand interferometry. The speckle pattern is derived from spatial frequencies less than (λ/D), whilst the interferometer fringes come from spatial frequencies much larger than (λ/D). Hence the interferometer fringe amplitude comes from the planet and not from the speckles, which only contribute the noise that was accounted for in the signal-to-noise calculation. Their assertion is, therefore, incorrect.

Angel *et al.*, in estimating noise sidelobes from their equation (1) for the quantity they call the gain, illuminate the fundamental issue. Equation (2) shows that the distribution of errors must be considered. The total power scattered from the main beam depends only upon ε, and if, for a given ε, l is small, the scattered power is distributed over a large angle, and the close-in sidelobe level will be low. That is why I chose l = 1 cm as the correlation length, giving a sidelobe suppression 1.6×10^3 greater than the example given by Angel *et al.*

On the other hand, if the only errors in the phase front were from large spatial frequencies, their objection would be well-taken. A beautifully polished surface, whose only errors in figure were on large scales, would be unsuitable for finding planets since all the noise sidelobes would be clustered close to the main beam. Thus, the example quoted by Angel *et al.*, who used l = 40 cm as the correlation length in their calculation, is concentrating on the wrong regime. Note that, in my discussion, a mirror diameter of 1.5 m was chosen for the receiving elements, because of the extensive industrial experience in making high-quality mirrors in this size range. Industrial suppliers, in informal inquiries, have indicated that achieving this precision of the element surfaces is an economic, not a technical, question.

In summary, it is far too early to choose between the optical interferometer and the monolithic infrared telescope for a planetary search, as both the optical and infrared approaches need more extensive study. The challenges are well-defined and there are no fundamental physical laws that forbid their solution.

B.F. BURKE

Department of Physics,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

1. Burke, B.F. *Nature* **322**, 340 – 341 (1986).
2. Angel, J.R.P., Cheng, A.Y.S. & Woolf, N.J. *Nature* **322**, 341 – 343 (1986).
3. Ruze, J. *Proc. Instit. elect. electron. Engng* **54**, 633–640 (1966).

ANGEL *et al.* reply—Burke has defined a mirror surface by equation (1) with l = 1 cm, showing that a relatively large r.m.s. error can be combined with very low sidelobes at 1/4 arc second radius. But this bears little relationship to the best optical surfaces currently made. In practice, errors are not confined to small scales of 'orange peel' and 'dog biscuit'. Zonal figuring, in particular, leaves circular waves on the crucial scale of 40 cm (λ/l = 1/4 arc s). For optical detection of a planet at this radius with apodized 1.5 m mirrors, sidelobes of 10–7 peak intensity need surface accuracy of about 1 Å r.m.s. on a scale of 40 cm, regardless of finer imperfections. By comparison, figuring errors on this scale are about 50 Å r.m.s. in the Hubble Space Telescope mirror. We would find the case for optical planetary detection more convincing if suggestions were given for dramatic breakthroughs in figuring, testing and support methods.

Does the use of an interferometer help, given an overwhelming need for high contrast rather than spatial resolution? For overcoming the photon noise of scattered light it offers at best no advantage over a single dish of the same area. The array gain cannot be better than the value of $D^2\theta^2/8\pi\delta^2$ given in our paper, where D is the diameter of the single dish that collects the same light. The character of the background is the same as the speckle-fringe convolution familiar to observers with ground-based optical interferometers[1], except at an appropriately reduced intensity level. As in the case for the single 8 m dish analysed in our paper, by morphology there is no way to distinguish a planet from the much brighter scattered starlight fringes. Only if the interferometer baseline is 40 m or more will a solar-type star at 4 pc be resolved and the scatterd light modulation be reduced. But an interferometer with so large a baseline, if arranged for good ultraviolet plane sampling, loses efficiency. Much of the planetary energy appears in parts of the diffraction pattern of low surface brightness[2], and the signal to noise ratio becomes much less than optimum.

We conclude that neither the optical spectrum nor interferometry are suited to this problem. The infrared telescope with filled aperture that we described has improved contrast and reduced scattered light for a given figuring precision, and can be accurately aligned by correcting for errors sensed at optical wavelengths.

ROGER ANGEL
ANDREW CHENG
NEVILLE WOOLF

Steward Observatory,
The University of Arizona,
Tuscon, Arizona 85721, USA

1. Hege, E.K., Beckers, J.M., Strittmatter, P.A. & McCarthy, D.W. *Appl. Opt.* **24**, 2565 (1985).
2. Meinel, A.B., Meinel, M.P. & Woolf, N.J. *Appl. Opt. opt. Engng* **10**, 150 (1983).

Announcing a major, new publication from The British Mycological Society

Covering all aspects of fungi and their uses in science, industry and conservation this exciting new journal will be launched in January 1987. Written mainly for the professional mycologist, **The Mycologist** will appear quarterly and is published by Cambridge University Press. Regular features will include mycotechnology, photographing fungi, identification of fungi, book reviews, developments in mycology and news of international mycological associations. **The Mycologist** is available on subscription for £6.00 ($12.00) per annum.

Order from Cambridge
University Press,
The Edinburgh Building,
Shaftesbury Road, Cambridge
CB2 2RU, England

Cambridge Journals

Reader Service No.2

NEW

THE SOFTWARE DIRECTORY FOR MOLECULAR BIOLOGISTS

Christopher Rawlings

This single volume brings together the scattered information on software available to molecular biologists.

The Software Directory compares and critically appraises specialized software for molecular biology, and also provides you with concise descriptions of commercial and business software applicable to laboratory work.

Each entry covers functional features, machine type, computer manufacturer, operating system, and author or company. The book also contains a thorough index of all the software listed, under seven different headings (the five above, plus package name and keyword)—so that you can always find the software that is right for your work, quickly and easily.

The all-new *Software Directory for Molecular Biologists* gives you accurate information on all known software in the field—commercial or non-commercial. You will save time and money with this invaluable guide.

300pp
$80.00
0-943818-37-0

April 1986
£40.00
0-333-39821-1

CONTENTS

Part 1—Descriptions of commercial and business software and their use in the laboratory. Spread sheets, text processing, database management systems, terminal emulators, graphics, financial planning, file transfer, integrated software systems.

Part 2—Directory of specialist software for molecular biologists

Part 3—Indexes and cross references of software listed.

Glossary of terms

Order on FREE APPROVAL by mail or phone. We accept all major credit cards.

US/Canada Mail to:
Stockton Press
15 East 26th St.
New York, N.Y. 10010
Call toll-free 800-221-2123
In N.Y., 212-481-1334

UK/Rest of World Mail to:
Elaine Coles, Globe Book Services,
FREEPOST
Brunel Road, Houndmills
Basingstoke, Hants RG21 2BR, England
Basingstoke (0256) 26222
FREEPOST applies to UK only

Postage & Packing: US/Canada: Added to invoice at cost; prepaid free. UK/Rest of World: £2.00 per order UK; £2.00 per book Continental Europe; £3.50 per book rest of world. UK libraries free.

"... the series should become a standard reference work."
Medical Laboratory Sciences

Monographs on Pathology of Laboratory Animals

Sponsored by the International Life Sciences Institute (ILSI)

Editor: **T. C. Jones,** Harvard Medical School

These monographs provide authoritative microscopic and ultrastructural descriptions of naturally occuring or experimentally induced neoplastic and nonneoplastic lesions in several species of laboratory animals. The identifying pathologic features are correlated with toxicologic, biologic and behavioral data as well as with data from other species including man.
The international editorial board critically reviews all manuscripts. The frequency of specific lesions in different geographic regions and laboratory environment are depicted graphically.
The volumes make use of an internationally accepted nomenclature to improve worldwide communication among scientists, industries and governmental agencies investigating the safety of foods, drugs and chemicals subject to animal tests.

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Heidelberger Platz 3, D-1000 Berlin 33 · 175 Fifth Ave., New York, NY 10010, USA · 28, Lurke Street, Bedford MK40 3HU, England · 26, rue des Carmes, F-75005 Paris 37-3, Hongo 3-chome, Bunkyo-ku, Tokyo 113, Japan

Urinary System

Editors: T. C. Jones, U. Mohr, R. D. Hunt

1986. 36 figures, 44 tables. XVIII, 405 pages.
Hard cover DM 265,–. ISBN 3-540-16591-6

As in other volumes of this series, the structural aspects of natural and induced disease are used to identify specific lesions of the urinary system. The biologic significance of each lesion in then explored in detail for the rat, mouse and hamster. The effects of genetic and environmental background are evaluated whenever possible, and differences are described between species and strains of laboratory animals.

Digestive System

Editors: T. C. Jones, U. Mohr, R. D. Hunt

1985. 352 figures, 24 tables. XVIII, 386 pages.
Hard cover DM 292,–. ISBN 3-540-15815-4

Respiratory System

Editors: T. C. Jones, U. Mohr, R. D. Hunt

1985. 279 figures, 20 tables. XV, 240 pages.
Hard cover DM 245,–. ISBN 3-540-13521-9

Endocrine System

Editors: T. C. Jones, U. Mohr, R. D. Hunt
Consultant on Comparative Endocrinology: **C. C. Capen**

1983. 346 figures, 26 tables. XV, 366 pages.
Hard cover DM 220,–. ISBN 3-540-11677-X

3568/4/1

Prices are subject to change without notice.

Springer

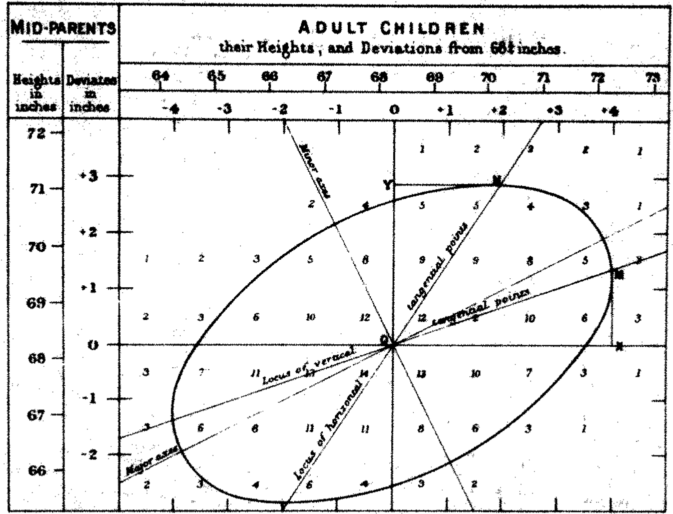
Reader Service No.3

Meaning from numbers

Ian Stewart

The Rise of Statistical Thinking 1820 – 1900. By Theodore M. Porter.
Princeton University Press: 1986. Pp.333. $35, £23.40.
The History of Statistics: The Measurement of Uncertainty Before 1900.
By Stephen M. Stigler. *Harvard University Press: 1986. Pp.410. $25, £21.25.*

THERE are many histories of mathematics. Statistics, a younger science, has been comparatively neglected. But now we have the happy conjunction of not one, but two fascinating books that describe the origins and early development of what is arguably *the* mathematical tool of this century. Both are scholarly yet readable; both cover the same period; and both are published by Ivy League presses. There the resemblance ends, and the books largely complement one another, the main overlap being the underlying historical reality.



Work of stature — Galton's 1886 plot of children's height against the average height of their parents, suitably smoothed to show the "series of concentric and similar ellipses" which brought together his ideas about regression in a single elegant mathematical package. (Reproduced from The History of Statistics.*)*

The Rise of Statistical Thinking avoids technicalities and concentrates on the flow of ideas between the natural and social sciences. It emphasizes the philosophical issues raised by novel statistical methods, and how they affected the subject's development. If it has a flaw it is that the technicalities have been smoothed away a little too much, so that one forgets that the founders of the subject were grappling with difficult mathematical, as well as social and philosophical issues.

The History of Statistics is stronger on the mathematical developments involved, especially the early history of probability theory, and is organized a little more explicitly around a selection of key figures. Mathematical detail is especially evident in the first half, where Stigler describes the contributions to probability theory of mathematicians such as Legendre, Laplace, Bernoulli and Gauss. This tends to give the impression that mathematicians have no part in human affairs except by virtue of their mathematics, whereas Laplace, for example, took a great interest in the implications of his work.

The most striking feature of the story is that both the "hard" and "soft" sciences played decisive roles in the development of statistics, and that important ideas and methods were transferred between them — in both directions. To illustrate this in microcosm, consider the "error law" or normal distribution. The name "error law" originated in the work of eighteenth-century astronomers and mathematicians, who, in trying to calculate the orbits of celestial bodies, were forced to take account of the effects of observational error. The idea was transferred into social science by Adolphe Quetelet. He was impressed, for example, by Laplace's probabilistic analysis of whether atmospheric pressure was subject to a regular cause of diurnal variation, and applied such analysis to everything he could think of: measurements of the human body, crime, marriage, suicide. By the early 1870s the idea had moved back into the physical sciences, with the development of statistical mechanics. In his book, Porter points out that both James Clerk Maxwell and Ludwig Boltzmann emphasized the success of the error curve in social science to justify importing it into physics. For example, in 1872 Boltzmann wrote:

> . . . the molecules of a body are indeed so numerous, and their movements so rapid, that nothing ever becomes perceptible to us except average values. The regularity of these mean values may be compared to the astonishing constancy of the average numbers furnished by statistics . . .

Quetelet was quick to draw general conclusions from the supposed constancy of mean values of social and anthropometric variables, and was the originator of the tantalizing notion of the "average man". He was influenced by the philosophy of Victor Cousin, and the idea of the *juste milieu* which simultaneously emphasized the supremacy of the bourgeoisie and embodied ideals of political moderation. To Quetelet, the "average man" was not just a mathematical abstraction, but a moral ideal. Others disagreed, and the applicability of the error law became highly controversial.

For Francis Galton, like Maxwell and Boltzmann, the error law was a means, not an end. Galton's interest in it arose in a curiously incidental way. A friend of his, William Spottiswoode, wrote a rather dreadful paper applying it to the alignment of mountain ranges. This paper drew Galton's attention to Quetelet's work.

Galton strongly opposed the interpretation of the law that anything exceptional was a flaw. To astronomers, variations in observations were errors, and the aim of the error law was to eliminate them. To Quetelet, deviation from the mean was a sign of moral or physical imperfection. To Galton, however, the variations themselves were the interesting phenomena, and the last thing he wanted to do was eliminate them. Galton, indeed, was the founder of eugenics, and he laid the ground for the "nature–nurture" debate. Always a forceful figure, in 1877 he campaigned to have Section F of the British Association — Statistics and Economic Science — expelled, because it wasn't mathematical enough. He also used the methods of social statisticians to debunk conventional religion: the most notorious of his writings on the subject was *Statistical Inquiries into the Efficacy of Prayer* in which he observed that sovereigns, whose lives were the object of regular prayerful appeal by entire nations, lived no longer than other members of the upper classes. Clergymen had a similar lifespan to physicians or attorneys. Insurance companies charged the same premiums to missionary ships and slave vessels.

Galton's most important contribution, to which Stigler devotes the best part of a chapter, was the idea of regression. Galton could not understand why the degree of variation of a characteristic did not increase from one generation to the next. Surely Quetelet's error law, with its implications of underlying homogeneity, *denied* the possibility of inheritance of characteristics? He eventually realized that a normally distributed mixture of normal distributions would combine to give one overall normal distribution. Each characteristic could follow its own error law. He built a mechanical device, the quincunx, in which lead shot bounced off pins (effectively performing a random walk), to demonstrate the principle. ▶

Comparing the distribution of the heights of children to those of their parents, Galton smoothed the data and obtained a dramatic example of what would now be called a multivariate normal distribution. This led him to the idea of regression lines and gave rise to the analysis of variance.

These ideas were taken up by Francis Ysidro Edgeworth, Karl Pearson and George Udny Yule, who between them took statistics into the twentieth century. According to Stigler, "Even in the varied world of nineteenth-century statisticians . . . Edgeworth was an anomaly". He was a classicist with no higher mathematical training who, while studying law, also undertook a massive programme of self-instruction in mathematics. From the beginning he sought links between mathematical ideas and areas such as psychology, economics and ethics. Soon he began to enquire into the statistical significance of irregularities in observations. At the jubilee meeting of the (Royal) Statistical Society he presented a paper on significance testing for the mean, and went on to develop an even more important idea: the correlation coefficient. Edgeworth changed the focus of statistics. Originally it had been about errors in observations. Now it was about the analysis of structure in data.

The nineteenth-century statisticians placed so much emphasis on the normal distribution that one might be forgiven for assuming that they recognized no other distribution curve. This is not entirely true, but they did expect observations of a single, homogeneous phenomenon to conform to the normal curve. Pearson challenged that viewpoint, and in his first publication on statistics he outlined an extensive theory of skew, or asymmetric, distributions. The full version ran to over a hundred printed pages and led to his election to the Royal Society. Pearson, whom Porter describes as a buccaneering type, was forever trying to grab new territory for statistics; in one memoir he applied his theories of skew variation to 14 examples, including the sizes of crabs, the American divorce rate and the number of petals on buttercups.

Yule was a student of Pearson, and when he arrived at Pearson's laboratory in 1893 he was probably expecting to study applied mathematics. But by 1895, says Stigler, "he was a statistician", and one showing signs of growing independence. Yule applied Pearson's skew curves to data on the frequency of paupers per poor-law union (combination of parishes for administrative purposes). Soon after he applied correlation methods to show that, contrary to assertions of Charles Booth in *The Aged Poor*, there was a relation between (in modern terms) the ratio of direct welfare to work-relief, and the incidence of poverty. In Yule's hands, Galton's eugenic idea of regression became a universal (though still controversial) tool for the analysis of cause and effect.

Thus, by 1900, the main foundation stones of classical probability theory and statistics had been laid. Both authors sensibly stop there, rather than plunging into the technical intricacies of the twentieth century. These are two important and substantial books which, between them, give a great deal of insight into the creation of a fundamental branch of science. In particular it is instructive to witness the way in which general mathematical concepts evolved from a morass of special problems about an enormous variety of topics, at a time when the most difficult question was to find the right questions to ask. □

Ian Stewart is in the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Northern waters

W.J. Gould & D.G. Masson

The Nordic Seas. Edited by Burton G. Hurdle. *Springer-Verlag: 1986. Pp.777. DM 198.*

"HAFVILLA" is a term used in the Norse sagas to denote getting totally lost at sea. In the scientific sense, *The Nordic Seas* might well provide us with a safeguard against such a fate.

The editor has collected together lengthy contributions from leading authorities in their respective fields, and has produced an impressive reference book on the oceanography (including studies of ice and climate), geology and geophysics of the seas lying between Norway and Greenland.

As is inevitable in such a compilation, there is a degree of duplication of material — for example between the chapters "Brief Overview of the Physical Oceanography" and "The Arctic Waters" — but this is a minor criticism. Over 200 pages are devoted to climatology, ice cover, oceanography (including sound speed) and tides, while the remaining 400 or so deal with geology and geophysics. A surprising omission is any mention of biology, because the subject is of interest to some physical oceanographers and the Nordic Seas are an area in which a great deal of work has been carried out on the relationships between physical and biological indicators and on the effect of physical parameters on fisheries.

The latter part of the volume deals with topography and geology. Much has been written on the geology of the area in recent years but this is the first truly comprehensive review, discussing everything from superficial sediments to sound velocity in the mantle. With its detailed subject index and comprehensive bibliography, it is an excellent "geological dictionary" and will be indispensable for those interested in the Nordic Seas and surrounding areas.

John S. Shelton

Outstanding feature — the 500-m high Ship Rock in New Mexico, a volcanic pipe which was exposed by erosion of its enclosing rock. The picture is taken from M.J. Selby's Earth's Changing Surface: An Introduction to Geomorphology, *published by Clarendon earlier this year. The book will be reviewed in* Nature's *annual textbook issue on 12 March 1987.*

The geological section has its disappointments, however. There appears to be little purpose in the straightforward description of bathymetry (Chaper 9), particularly because it is largely repeated and explained in terms of the underlying geology in the first part of the following chapter. Chapters 10 and 11 are simply too long; much of the considerable volume of data from areas peripheral to or analogous to the Nordic Seas could probably have been omitted without detriment to the overall content. Too much emphasis has also been placed on explanation of even the simplest geological terms — surely inappropriate in a work intended primarily for experts in the field. Finally, although it is otherwise well produced the book suffers from poor reproduction of many line drawings, to the extent that the chapter on bathymetry loses almost all of its impact because contour lines on the charts are invisible and only the annotation remains.

But the merits of the volume greatly outweigh the flaws. *The Nordic Seas* will find its way into the libraries of a wide range of scientific establishments and will be much referred to. It is as up to date as could possibly be expected and will remain as a key document for many years. □

W.J. Gould and D.G. Masson are at the Institute of Oceanographic Sciences, Brook Road, Wormley, Godalming, Surrey GU8 5UB, UK.

Parasitic puzzles

Sydney Cohen

Living Together: The Biology of Animal Parasitism. By William Trager. *Plenum: 1986. Pp.467. $71.40.*

William Trager has been a dedicated student of animal parasitism throughout his 50 years on the staff of the Rockefeller Institute in New York. During most of that time the subject was a relatively neglected one. The simpler prokaryote parasites and their host interactions are more amenable to study and received much greater attention. The recent surge of interest in animal parasitism reflects an increased awareness of the global impact of animal parasites as human pathogens, but stems also from several technological advances which have transformed the feasibility of carrying out experiments in the subject. Among such advances is Trager's own most notable scientific achievement, with Jensen in 1976, of cultivating for the first time the human malaria parasite, *Plasmodium falciparum*.

Animal parasites are extraordinarily complex and adaptable organisms. During successive life cycles in different host species, individual parasites may assume forms so disparate as to have been originally regarded as separate species. Alternating life cycles may also involve abrupt changes in metabolism, such as the switch from the glycerophosphate oxidase system of the bloodstream form of African trypanosomes to the cytochrome-mediated respiration of the insect form which is non-infective for its mammalian host. The study of animal parasitism is therefore basically an attempt to understand the sequential control of gene expression which generates alternative parasite forms adapted to recognition, penetration, growth and reproduction in hosts as different as (for example) the sandfly and man. In addition, the student of animal parasitism must attempt to delineate the complex reactions of the host to invasive challenge.

In confronting this daunting task, Trager deals first with a series of fundamental biological events. He starts with parasite recognition and entry into the host, and their attachment to and penetration of specific host tissues and cells. There follows a discussion of the interface between host and parasite and the molecular exchanges across these surfaces that are vital for parasite survival. Host reactions to parasite invasion are considered in relation to innate immunity, the important and often mysterious role of the spleen, acquired immunity and the pathogenesis of parasitic disease. In the latter part of the book, phenomena of parasitism are discussed in relation to a selection of specific protozoan and helminthic parasites of human beings and other vertebrates that are currently the centre of research work. The closing chapters deal briefly with symbiosis and with the methods traditionally used for the control of parasitic diseases, namely, chemotherapy and ecological methods such as vector control.

Throughout the text, Trager draws attention to the innumerable questions raised by the phenomena he describes. In a few instances these can be satisfactorily accounted for in molecular terms. For example, we at last have an answer to the long-standing puzzle of why West Africans and their descendants in the New World are totally refractory to infection with *Plasmodium vivax* and yet are susceptible to the other plasmodial species that infect mankind. Also, the means whereby the sickle-cell gene engenders resistance to infection by malignant tertian malaria is understood in detail. A great deal is also known about the structural and genetic basis of antigenic variation in African trypanosomes which mediates their evasion of potentially lethal immune attack by the mammalian host. For the most part, however, the biological phenomena described by Dr Trager with arresting simplicity provide a Pandora's box of puzzles which will stimulate and challenge future generations of molecular pathologists. The answers — when they come — will illuminate not only the mechanisms of animal parasitism but also our understanding of biological differentiation.

In his introduction to the book, Joshua Lederberg remarks that descriptions of the lives and adaptations of animal parasites are "a veritable Arabian nights of narrative, not of the human imagination, but of Nature's". The subject is indeed a fascinating one, and also is of great consequence in human health and welfare. This book, distilled from a lifetime's work, and set out in uncluttered style with abundant illustrations, will be read with profit and enjoyment by a wide audience including both aspiring and established molecular parasitologists. During the past decade many imaginative initiatives have transformed the study of animal parasitism by bringing together research workers from a variety of disciplines. Dr Trager's book will give further impetus to recruitment into this engrossing and important area of research. ☐

Sydney Cohen is Emeritus Professor of Chemical Pathology in the United Medical and Dental Schools of Guy's and St Thomas's Hospitals, London SE1 7EH, UK.

Unsure accuracy

Alan J. Charig

The Riddle of the Dinosaur. By John Noble Wilford. *Faber & Faber/Knopf: 1986. Pp.304. £14.95, $22.95.*

One might assume from the title that this is just another book on dinosaurs; after all, there are already so many on the market that thinking up a new title must always be a problem. But the exact wording hints strongly at the distinctive nature of this excellent volume. For it is not a descriptive work, listing the various genera of dinosaur and giving an account of the anatomy, physiology and habits of each, nor is it even concerned with their evolutionary history; rather does it deal with the history of dinosaur discovery, with dinosaur research and with the many controversies (both past and present) that bedevil the life of the dinosaurologist.

The text has been well researched and is factually correct. It is also stylishly written, and is *interesting*, even to the professional dinosaur worker. Most important of all, it reports on the controversies in a balanced, impartial manner, often coming to no positive conclusions but leaving the reader sitting on the fence — which, in our present state of knowledge, is in many cases the proper place to be.

The author has based much of his book on interviews and discussions with well-known scientists active in the dinosaur field. Since he himself is an American, a Pulitzer Prize-winning science correspondent for *The New York Times*, it is understandable that those people he chose to talk to were all from the United States; the book therefore gives the impression that dinosaurology today is an almost exclusively American subject. His historical perspective nevertheless embraces a fair amount of material from outside North America; it includes — as it must — the story of the first discovery of dinosaur remains, in England, and the initial perception, by Mantell and subsequently Cuvier, that reptiles of enormous size and (as originally conceived) herbivorous diet had existed in Mesozoic times.

My only complaints are to do with the illustrations, which are extremely few. One of them, the family tree on p. 189, is familiar; I recognize it from another volume, although its source is nowhere acknowledged. In the middle of the book are seven colour restorations by Douglas Henderson; his dinosaurs look wooden and his pictures as a whole have neither scientific nor aesthetic appeal. Moreover three of the seven are given the wrong date.

That notwithstanding, here is a book that no dinosaur addict — be he a professional or one of those innumerable keen amateurs — can afford to be without. ☐

Alan J. Charig is in the Department of Palaeontology, British Museum (Natural History), Cromwell Road, London SW7 5BD, UK.

Is the reductionist beyond belief?

William Paton

God and the New Biology. By Arthur Peacocke. *Dent: 1986. Pp.198. £10.95. To be published in the United States by Harper & Row.*

"THE brain secretes thought as the liver secretes bile; and you are all just bundles of conditioned reflexes". With this ending to his lectures on the alimentary tract, my physiology professor, the late John Mellanby, leant on his elbow on the lectern, and beamed at us. Although he seemed a pretty implausible reflex-bundle himself, that, 50 years ago, was my first introduction to reductionism. A.J. Ayer's *Language, Truth and Logic* was also giving logical positivism a brisk run. An advanced cleric of the day, L.W. Grensted, was said to believe that "There is no God, men are monkeys and there may be spooks". To the eye today, it was all good clean fun.

Faith, reductionism, the "enfant terrible" and the "advanced cleric" are still with us — but the antitheses are less sharp. The church has learnt to face the criticisms of rigorous scholarship, and science that the ethical problems it can create need ethical principle not of its making for their resolution. Leaders of science and religion are now less prone to pronounce at each other's expense.

THE MANAGEMENT OF AIDS PATIENTS

Edited by DAVID MILLER, JONATHAN WEBER and JOHN GREEN

The Management of AIDS Patients is the first comprehensive guide to the practical clinical management of patients with AIDS or HTLV III infections. The book avoids the sensational aspects of the disease, offering solid advice and information for all people involved in patient care.

The editors and many of the contributors come from St Mary's Hospital, London one of the foremost centres in Britain for the treatment of AIDS patients. The knowledge and experience of these experts has been combined to provide a much-needed book for doctors, nurses, dentists and indeed for health-care professionals everywhere.

1986

200pp Hardback 0 333 40465 3 £30.00
Paperback 0 333 40466 1 £10.95

Please order this title from your bookseller – or in case of difficulty from Judy Chappell, Macmillan Press, Houndmills, Basingstoke, Hants RG21 2XS.

M MACMILLAN PRESS

Dr Peacocke's book on biological reductionism is admirably consonant with this atmosphere. He is well-qualified; a physical biochemist and tutor in Oxford, who took orders, became a college chaplain in Cambridge and is now director of the Ian Ramsey Centre in Oxford, an institution recently founded for the study of ethical problems in science and medicine. The book pulls together his thinking over a number of years.

His conclusion is best given in his own words (though they are more precise than colourful):

> It *is* possible for higher level concepts and theories... to be non-reducible to lower level concepts and theories, that is, they can be autonomous. At the same time one has to recognise the applicability of the lower level concepts and theories (for example, those of physics and chemistry) to the component units of more complex entities and their validity when referred to that lower level. That is, with reference to biology, it is possible to be anti-reductionist without being a vitalist. Higher level concepts and theories often refer to genuine aspects of reality at their own level of operation and we have to eschew any assumptions that only the so-called fundamental particles of modern physics are "really real". To do so is indeed a kind of "fundamentalism" of a pseudo-scientific kind that is as much to be avoided in impartial enquiry as its religious namesake.

The book is a critical review of a rather extensive literature on and around the "reducing" implications of modern biology, molecular biology, genetics, evolution and sociobiology (but, deliberately, not neuroscience).

Dr Peacocke finds a genuine middle ground. The reductionist refuses to accept any "extra" (for example, entelechy or vital principle) beyond the laws of natural science in order to explain natural phenomena. The anti-reductionist finds it inconceivable that the full variety of biology and human life can be so explained without it. But molecular biology and genetics do bring one striking gift: the breadth of their explanatory power shows how, from simple elements, systems of astonishing complexity and richness can arise. Randomness presents itself not as some irrational intrusion, but as a source of creativity. And it seems that in these new systems, there is a kind of autonomy. Social behaviour, for instance, may indeed depend on the existence of intermolecular forces, but it is to be discussed and experimented on in terms of its own concepts rather than, say, that of the Van der Waals equation. So, too, the B Minor Mass is a "mere" assembly of diatonic tones — yet we have to understand and discuss it in musical, not acoustic language. If, still on the threshold of knowledge, we can already see that in the primary chemical elements there lies the potential for the synthesis of proteins and enzymes, and cell replication, cell-constituent assembly, speciation, learning and association, what may there not be still to be revealed?

Like other intermediate positions, it may not satisfy. Does it leave, for the Christian, room enough for a God with a loving personal relation to his creation? Does it provide, for the reductionist, too much room for self-deceiving religious aspirations to creep in? Perhaps it is not so much the possible existence and human awareness of some higher being that can most worry a scientist, but the idea of that being interfering with the natural course of events, with "natural law". Yet the question is not much different from the age-old one of whether our own day-to-day decisions (seemingly "free") actually modify events or not. If our belief in the one influence is valid and compatible with natural law, why not the other?

Dr Peacocke, though unequivocally Christian, is not, I think, pressing his views. His posture is like that of Bishop Butler in the *Analogy of Religion*, published almost exactly 250 years ago:

> It is come, I know not how, to be taken for granted, by many Persons, that Christianity is not so much as a Subject of Enquiry; but that it is, now at length, discovered to be fictitious. And accordingly they treat it, as if, in the present Age, this were an agreed Point, among all People of Discernment; and nothing remained, but to set it up as a principal Subject of Mirth and Ridicule, as it were, by way of Reprisals, for having so long interrupted the Pleasures of the World. On the contrary, thus much, at least, will be here found, not taken for granted, but proved, that any reasonable Man, who will thoroughly consider the Matter, may be as much assured, as he is of his own Being, that, however, it is not so clear a Case, that there is nothing in it.

For those who do not like to see people bludgeoned into belief, and who are also interested in the work of recent years by both a theologian and a scientist, Dr Peacocke's eirenic writing will be much welcomed. ☐

Sir William Paton, 13 Staverton Road, Oxford OX2 6XH, UK, is Emeritus Professor of Pharmacology at the University of Oxford.

New in paperback

- *The Birth of Particle Physics* edited by Laurie M. Brown and Lillian Hoddeson (reviewed in *Nature* **308**, 383; 1984). Publisher is Cambridge University Press, price is £12.95, $18.95.
- *The Ecological Web: More on the Distribution and Abundance of Animals* by H.G. Andrewartha and L.C. Birch (reviewed in *Nature* **315**, 81; 1985). Publisher is University of Chicago Press, price is $19.95, £16.95.
- *Measuring the Universe: Cosmic Dimensions from Aristarchus to Halley* by Albert Van Helden (reviewed in *Nature* **316**, 492; 1985). Publisher is University of Chicago Press, price is $8.95, £7.50.

Wiley

FUNCTIONS OF THE PROTEOGLYCANS

Ciba Foundation Symposium 124

Chairman: **V. Hascall,** *National Institute of Dental Research, NIH, Bethesda*

Proteoglycans are macromolecules containing glycosaminoglycan chains covalently bound to protein cores. In this symposium workers from a range of disciplines (biochemistry, cell biology, pathology and developmental biology among others) focus on the structure and functions of the many types of proteoglycan found in mammalian tissues.

Series: Ciba Foundation Symposium

0471 91090 2 310pp November'86 £27.50

ANTIDEPRESSANTS AND RECEPTOR FUNCTION

Ciba Foundation Symposium 123

by **D. Murphy,** *National Institute of Mental Health, Bethesda, USA*

Depression is a common and often debilitating affective disorder. Attempts to develop effective antidepressants have a long history, but many questions remain about the mechanisms of action of such treatments and about the aetiology and pathophysiology of depression itself. Pharmacological studies, aimed ultimately at elucidating the neurochemical basis of depression and of promoting new therapeutic approaches, provide the focus of this symposium volume.

Series: Ciba Foundation Symposium

0471 91089 9 304pp November'86 £27.50

ULTRATRACE ANALYSIS OF PHARMACEUTICALS AND OTHER COMPOUNDS OF INTEREST

by **S. Ahuja**

This ready reference describes analyses performed in a variety of areas and how they are instrumental in solving some of the complex scientific problems relative to drugs, food safety, pharmaceuticals, air, food, water and the environment.

Series: Chemical Analysis

0471 82673 1 400pp November'86 £57.50

MASS SPECTROMETRY IN BIOMEDICAL RESEARCH

edited by **S.J. Gaskell,** *Tenovus Institute for Cancer Research, University of Wales College of Medicine, Cardiff, UK*

This book provides a state-of-the-art coverage of the dramatic extension in the application of mass spectrometry (MS) to biomedical research that has taken place in the last few years. Dr Gaskell has brought together a large number of eminent and innovative researchers in MS, whose major motivation is the desire to solve problems in biomedical research.

0471 91045 7 508pp November'86 £38.50

RUSSIAN FOR THE SCIENTIST AND MATHEMATICIAN

by **C.A. Croxton,** *Department of Mathematics, University of Newcastle, New South Wales, Australia*

This book is intended to enable scientists and mathematicians to produce their own translations of Russian technical literature, which is otherwise slow and expensive to translate.

0471 90260 8 226pp April'84 £21.45

John Wiley & Sons Ltd.
Baffins Lane, Chichester
Sussex PO19 1UD, England

Reader Service No.1

CALL FOR PAPERS:

JOURNAL OF EVOLUTIONARY BIOLOGY

First Issue: January, 1988
(Will be published in October 1987)

The journal will publish original research articles in English spanning, on the one hand, *theory, botany, and zoology,* and on the other, *ecology, genetics, development, and paleontology.* It is an international journal welcoming contributions from all countries. For format, please see a recent issue of the journal Evolution.

Manuscripts may be submitted either to the
Managing Editor,
Prof. S.C. Stearns,
Zoologisches Institut,
Rheinsprung 9,
CH-4051 Basel, Switzerland,
or to a member of the Editorial Board:

Richard Abbott (St. Andrews)
Graham Bell (McGill)
William Chaloner (London)
Gary Freeman (Austin)
Tony Hallam (Birmingham)
Suresh Jayakar (Pavia)
Shoichi Kawano (Kyoto)
Eduard Linsenmair (Wurzburg)
Elisabeth Lord (Riverside)
Karl Niklas (Ithaca)
Linda Partridge (Edinburgh)
Rudolf Raff (Bloomington)
Joe Travis (Tallahassee)
Pere Alberch (Harvard)
Michael Bulmer (Oxford)
Hugh Dingle (Davis)
Pierre-Henri Gouyon (Montpellier)
Erkki Haukioja (Turku)
Gerdien de Jong (Utrecht)
Marty Lechowicz (McGill)
Volker Loeschke (Aarhus)
Hans Metz (Leiden)
Arie van Noordwijk (Basel)
Wolf Reif (Tubingen)
Valerio Sbordoni (Rome)
Gunter Wagner (Vienna)

First Annual Meeting of the EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

26–30 August 1987 Basel, Switzerland

Plenary lectures by distinguished speakers.

Workshops and symposia with contributed talks.

A business meeting to approve bylaws, elect officers, and establish the Society's relationship to the Journal of Evolutionary Biology.

Please address suggestions and requests for information to: Dr. A.J. van Noordwijk, Zoologisches Institut, Rheinsprung 9, CH-4051 Basel, Switzerland.

Birkhäuser Verlag
Basel Boston Stuttgart

Reader Service No.61

CONVENIENT·INFORMATIVE·TIMELY·PROVOCATIVE

FOUR DYNAMITE REASONS WHY YOU SHOULD SUBSCRIBE TO BIO/TECHNOLOGY NOW . . .

1. You'll save time. Because you don't have time to read every journal in the field, BIO/TECHNOLOGY is written and edited to give you all the information you need in ONE convenient source. BIO/TECHNOLOGY is the only journal to give you news. Features. Original research. Product evaluations. Industrial applications. Job openings. Book and product reviews. And more. If it's important to your work, you'll find it between the covers of BIO/TECHNOLOGY.

2. You'll learn new ways to improve your on-the-job skills and run a more successful operation. Each month BIO/TECHNOLOGY gives you the scoop on invaluable start-up information . . . the latest instrumentation . . . marketing breakthroughs . . . funding sources . . . new technologies . . . patents and licences. . . profiles on how others are avoiding the pitfalls and reaping the rewards. In fact, just one practical tip from BIO/TECHNOLOGY can be worth many times the cost of your subscription.

3. You'll be up-to-date on all major industry developments. With correspondents all over the world, BIO/TECHNOLOGY delivers the inside story on who's doing what, where. You'll get the latest word on changing government regulations and tax law, and how they affect your work.

Plus you'll read original research papers by bench scientists, engineers, and biologists that will help keep you in the forefront of every new discovery worth knowing about.

4. You'll get new ideas you can put to use immediately. BIO/TECHNOLOGY lets you rub elbows with eminent scientists, tour new facilities and learn how new processes work in every field from genetic engineering to molecular biology. Each month, BIO/TECHNOLOGY helps spark ideas for your own research, for new industrial applications and for getting along in the increasingly competitive financial arena.

If you're not already subscribing to BIO/TECHNOLOGY, don't delay the moment any longer - return the coupon below and we will send you twelve issues. You risk nothing. If for any reason BIO/TECHNOLOGY doesn't live up to your highest expectations, just let us know and you'll receive a full refund on all unmailed copies. Remember too, that your subscription may be tax deductible.

TO START RECEIVING YOUR SUBSCRIPTION, FILL OUT THE COUPON BELOW AND RETURN WITH PAYMENT TO:

U.S.A.: BIO/TECHNOLOGY
Subscription Order Dept., P.O. Box 1543
Neptune, New Jersey 07754-1543 U.S.A.

To speed up your subscription in the U.S.A., call toll free 1-800-524-0328.

ENGLAND: **BIO/TECHNOLOGY**
Subscription Order Dept., Macmillan Journals, Ltd.
Brunel Road, Houndmills, Basingstoke
Hampshire RG21 2XS, England

☐ YES! Please send me one year of BIO/TECHNOLOGY (12 issues). I understand that I may cancel my subscription at any time and receive a full refund on all unmailed issues.

☐ U.S.A.-$59.00* ☐ Australia-A$130**
☐ U.K.-£70** ☐ Elsewhere-U.S.$112**

Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan Tel: (03) 292 3755

Note: *Represents 24% off Basic Subscription Price of U.S. $78.00
**All other prices quoted are "Air Speed"

☐ My payment is enclosed at the price checked.
☐ Please charge my credit card account:
☐ American Express ☐ Mastercard ☐ Visa

Card No.: ____________ Exp. Date: ________
Signature: ____________
Name: ____________
Company: ____________
Address: ____________
City/State/Postal Code: ____________
Country: ____________

7AUS

NATURE VOL. 324 11 DECEMBER 1986 COMMENTARY 523

The music of digital computers

Philip Campbell

More and more composers are relishing the unprecendented freedom provided by digital processors to manipulate both internal and external characteristics of sounds.

AMONG the most laughable events in the history of music—according to the cartoonists of the time at least (Fig. 1)— were the concerts of the Italian futurists during the first decades of this century. Depicted thus as sonic anarchists, contemporary photographs and manifestos nevertheless indicate the high seriousness with which, with the help of their 'noise-intoners' ('*intonarumori*'), the futurists fused the sounds of everyday machinery and mechanically synthesized noise into no-doubt outrageous cacophony. Though that artistic movement proved still-born, more recent figures, driven by a similar determination to escape what they see as the shackles of the past, have provided the composer with an instrument of more profound potential—the digital computer, capable of dissecting, transforming or synthesizing sounds according to a limitless variety of prescriptions.

The association of the computer with music goes back several decades. Composers such as Milton Babbit have used it since the 1940s to create ultra-rigorous twelve-tone music while, at another extreme, Iannis Xenakis has used it to inject a more statistical component into musical design[1]. Others have used computers for musical analysis and for modelling aspects of musical perception (see, for example, ref. 2).

But it is only now that the capacity of the digital computer is beginning to approach that required to make it a truly flexible tool for the creative musician—not a crutch by which deficient composers can save themselves work but a doorway to new paths of musical imagination, in particular those that encompass the microstructures of sounds.

In this article I shall concentrate on two particularly stimulating and poetic realisations of the computer's potential. Both, as it happens, were produced at what is currently the mecca of computer-assisted music, the Institut de Recherche et de Coordination Acoustic/Musique (IRCAM)—a subterranean establishment adjacent to the Pompidou Centre in Paris. The first, *Mortuos Plango, Vivos Voco* by the English composer Jonathan Harvey, is a composition on tape, about ten minutes long, which manipulates and intertwines only two sounds—a bell and a singing voice. The second, *Répons* by Pierre Boulez, is yet to be completed but, through several preliminary exposures, has already established itself as a benchmark in the use of computers to transform the sounds of musical instruments during a live performance.

Consider the scale of the problem. Beethoven's ninth symphony reaches our ears in the form of a complex pressure signal lasting about one hour. The frequency range of the best human ears extends upwards of 16 kHz; to satisfy them, a digital device wishing to reproduce the aural magic of, say, the Berlin Philharmonic must be able to synthesize or reproduce waveforms up to such frequencies—we can safely set the upper limit at 20 kHz. The principle of digital operation involves breaking up the continuous signal into a series of discrete samples and, because one has to have at least two samples per oscillation to reproduce a wave at a given frequency, we require a sampling rate of 40 kHz. Furthermore, the amplitude of the waveform as it varies from sample to sample must be resolved with good precision, otherwise noise ('quantization errors') will be introduced. Sixteen bits are usually considered as an adequate representation for each sample ($2^{16} = 65536$).

Thus simply to store Beethoven's ninth symphony in a digital memory requires about 14×10^9 bits; to synthesize it from scratch would represent a truly gargantuan feat of calculation (and perhaps a sterile one at that). For any instrument one might require a digital oscillator for each harmonic component of the waveform, each of a relative amplitude that, as it happens, varies according to the pitch and loudness of the note being played. Furthermore, to sound natural each harmonic component has to have the appropriate relative phase together with a small degree of random variability. Finally, the resulting tone has its own 'global' envelope determined by the composer's score, variable loudness as in a crescendo, for example, or perhaps a long sustained note of rapidly re-iterated attacks, a tremolando.

It is clear that to provide a flexible digital instrument for the composer requires truly formidable computer power. Necessarily, both of the pieces to be discussed here make do with considerably

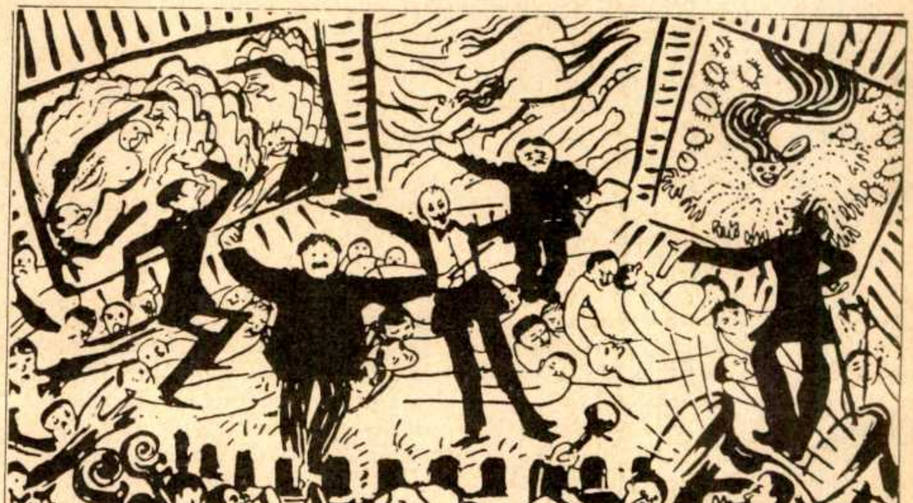


Fig. 1 A futuristic evening (Boccioni, 1911).

less than the full real-time synthesis described above. The tape of *Mortuos* was painstakingly accumulated over months of work in a studio (there is no musical score—to perform the piece one simply plays the tape). The live transformations of *Répons* primarily involve relatively simple arithmetic operations on incoming signals. Of course the musical power of such techniques springs not from their degree of sophistication but from the sureness of judgement with which they are married to strong ideas.

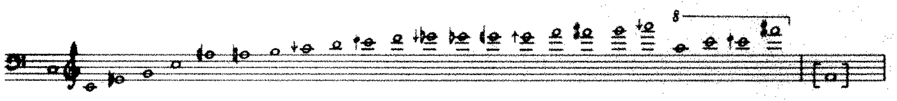


Fig. 2 Spectrum of the tenor C bell of Winchester cathedral in musical notation. The final F in parentheses is a fictitious harmonic, apparent only to the human ear deceived by upper harmonics[4]. Key: 𝄲 sharpened (raised) by a quartertone: ♯ sharpened by a semitone; 𝄳 flattened (lowered) by a quartertone; ♭ flattened by a semitone; ↑ adjusted by less than a quartertone.

The music of *Mortuos*

In the compact Norman tower of Winchester Cathedral can be found a set of seventeenth century bells, on the largest of which—the tenor C bell—are inscribed the words "*HORAS AVOLANTES NUMERO, MORTUOS PLANGO: VIVOS AD PRECES VOCO*" ("I enumerate the fleeing hours, I mourn the dead, I call the living to prayer"). Two sounds provide the basic elements from which *Mortuos* is constructed: the peal of this bell and the treble voice of the composer's son singing these words.

The publishers of *Nature* are, alas, insufficiently generous to provide the reader with a complimentary recording of *Mortuos*. Moreover the only commercial recording currently available[3] is unlikely to be found on the shelves of any but the most enlightened record stores (of which there appear to be none in London, for example). In the face of such obstacles, the following description of the music will have to suffice.

Any attentive reader of newspapers will have noticed the old trick by which his or her attention is sought—the classic first paragraph in which the essence of a story is presented with maximum impact. Harvey achieves precisely this, opening *Mortuos* with a vigorous carillon of bells which gradually declines in intensity until only two or three slowly pealing bells are left, over which the boy intones part of the inscription from which the title is drawn. Both bell and boy sounds, here and throughout the piece, are digitally synthesized or reproduced.

The significance of this carillon is greater than might be thought at a first hearing, when it is likely to appear as a random clustering of bell sounds. As it happens the composer is here representing in macroscopic terms the microstructure of the tones of the C bell. The harmonics produced following a bell stroke, which Harvey analysed (4) by a standard fast Fourier Transform, are shown in Fig. 2. During the reverberation these harmonics decay in descending order. This microprocess is writ large in the opening carillon in that each of the bells that simultaneously sound is repeatedly playing one of these harmonics; the higher the pitch of the bell, the more rapidly it repeats its stroke and the sooner it ceases, giving rise to the cumulative and compelling effect described.

Each of these 'bells' was synthesized with a spectrum identical, in relative terms, to that in Fig. 2, being simply a reproduction of the C bell transposed upwards in pitch. The reverberating tones of the C bell were digitally sampled and stored in a 'wave-table' within the computer memory which could itself be sampled at varying speeds, both backwards and forwards. This facility allowed Harvey not only to generate the opening carillon but also, later in the piece, to manipulate and overlay the pattern of decaying harmonics to yield subtle rhythmic effects or, by simply reversing the order of fading, to 'turn the bell inside out'[4].

Having caught his listener's attention with this declamatory exposure of his basic materials, Harvey proceeds to unfold the only element of the composition which might be called a 'theme': the slow succession of notes drawn from the C bell spectrum, shown in Fig. 3. Each note is played by bell tones 'faded' in and out so that successive notes overlap, their combined spectra providing a high frequency sheen above the notes. Such effects of timbre, together with harmonic juxtapositions such as that between an A♭ in the spectrum of the note F and the following A♮ of the "melody", all in the context of the quiet simplicity of the theme itself, combine to yield a memorable and beautiful statement.

This gesture opens the first of the eight principal sections of the piece, each of which begins with one of these thematic notes struck by a bell: within each section that note provides a basis for computer-aided musical development. Thus the entire piece adds up to an elaboration of the theme, all built with the materials presented at the outset.

In one sense the music is tonal (in contrast to the atonal style prevalent in music written over the last two or three decades), in that a pitch from which the material springs is fairly explicit at any point in the composition. The intervals within the primary bell spectrum, particularly the fourth, fifth and octave, pervade the piece—though a unique and essential flavour is added by those notes most removed from the well-tempered scale. But at no point is the writing harmonic in a progressive fashion (as in the classical style, where one chord leads to another via movement of the bass). In that sense the piece is static—a contemplation, if you like, of the inner life of certain sounds. Nevertheless, the flow of Harvey's invention creates a momentum of its own.

The techniques by which each of the thematic notes is elaborated within each section represent something of a showcase for the potential of the computer for sound sculpture. For instance the second section (based on G) leads to the third (F) by an unprecedentedly literal means of harmonic preparation: from a chord comprised of a combination of boy and bell tones, some of the notes suddenly slide upwards and/or downwards in pitch (an effect known as a glissando) to arrive at a new set of notes forming part of the spectrum of the F bell stroke which follows. Here it is the synthesized voices of the boy which execute the glissandi whereas, in the third section, it is the sine wave components of the bell that do the work in what turns out to be a study of glissando manipulation.

In the fifth section such simultaneous glissandi, executed in clusters of voices with added reverberation, provide a spectacular quasi-choral backcloth. Harvey organises and overlays the glissandi so as to generate a sense of chordal progression behind fragments of text sung solo in the foreground.

Each section of *Mortuos* exhibits at least one distinct facet of the digital computer's ability to render malleable the internal as well as the external characteristics of sounds, varying the juxtaposition of the harmonic components such that bell and boy at times merge imperceptibly into one another—with the bonus of spatial effects added by means of the eight tracks of the tape which can be independently routed to different speakers.

The eighth section provides a final example. Here the synthesized C bell peals with slow repetition while a choir of boys

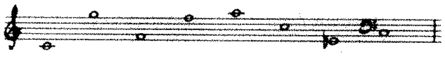


Fig. 3 The eight primary pitches from the bell spectrum which define both the 'theme' and the principal sections of *Mortuos*.

periodically 'tolls' a dissonant note cluster derived from the C bell's spectrum. But the harmonics of the bell are differentiated by the manner in which their loudness is varied. Some are sustained over many strokes while the majority follow the pattern of impact and subsequent decay, but with added slow variation within the overall decay envelope. This simulates the effect of the swing of a bell—for a static listener, the perceived spectrum varies cyclically with the changing orientation. And in an inspired final gesture, Harvey removes the initial sounds of the impacts, maintaining the 'swinging' effect within the fading peals. It is as if the air itself were echoing the vanished bell.

Synthesis in *Mortuos*

The techniques used by Harvey (together with his technical assistant at IRCAM, Stanley Haynes) represent an appreciable proportion of those currently available for digital synthesis. (For an excellent overview see ref. 5 or refs. 6 and 7). Once the harmonic components of the C bell had been determined, its sound could be synthesized by combining sine tones of appropriate frequencies and relative amplitudes and phases—a technique known in the trade as additive synthesis. Another approach adopted, known as fixed waveform synthesis, was to store the digitally sampled sound in successive memory locations—the wave table technique already mentioned. Finally subtractive synthesis, in which a broad-frequency-band source is subjected to appropriate filtering, was used to reproduce the boy's voice.

Additive synthesis. The complexities of building up the highly varied sounds of musical instruments from their components have already been mentioned. Bell sounds, at least in their natural form, are somewhat simpler to reproduce. By means of the Music V software at IRCAM[8,9], Harvey was able to generate both the natural bell sounds and the control of individual harmonics described above. Additive synthesis also permitted interpolation between two tone qualities or timbres at a given pitch; thus by appropriate variation of the harmonics a bell sound could gradually be transformed into a treble voice or vice-versa. Furthermore, by sending different harmonics to different speakers the composer was able to conjure up the illusion of being inside the bell.

Fixed waveform synthesis. By storing sound in a wavetable, one loses the ability to manipulate its harmonic components but gains significantly in the ease with which the sound as a whole can be transformed. For example one can easily transpose the sound to a higher pitch by increasing the effective rate at which the waveform is sampled. For a given rate of access to the memory, one can double the frequency (transpose up an octave) by sampling every other location along the wavetable. Of course this procedure distorts the waveform, making its envelope less smooth and thereby introducing noise, but sampling rates are typically high enough for this not to matter. Another problem occurs if there are harmonic components present whose frequency is higher than half the effective sampling frequency. Such components will be insufficiently sampled and will lead to "aliasing"—the generation of spurious signals at lower frequencies.

Subtractive synthesis. One of the longstanding goals of sound synthesis is that of mimicking the human voice—the ultimate user-friendly computer will of course use speech synthesis at a level indistinguishable from that of humans. The approach adopted by most working in this area is that which our species has evolved: a sound source subsequently modified by a series of resonators that provide individual characteristics as well as common elements of language such as vowel sounds. These manifest themselves as distinctive features in the spectrum superimposed on the fundamental tone of the speaking or singing voice. The source in the human case is a double reed (the vocal chords) which modulates a stream of air travelling from the lungs through the trachea, a tube of continuously varying diameter which in turn feeds several cavities including the mouth and nose. The lips, tongue, palate and glottis can introduce turbulence or sudden perturbations into the flow to produce 'fricative' and 'plosive' sounds respectively. The aim of subtractive synthesis is to mimic this system with appropriate sound sources and filters.

To reproduce his son's voice for *Mortuos*, Harvey used the CHANT algorithm developed at IRCAM by Xavier Rodet and others[10]. This digital approach generates fricative and plosive sounds by means of a flat-spectrum source coupled to appropriate filters, generating a series of pulses spaced according to the desired fundamental frequency. This output is then fed through a more complex set of filters to reproduce the effects of the resonant trachea, larynx, mouth and other cavities. (At the time that *Mortuos* was written the CHANT algorithm could not simulate consonants, which were therefore digitized directly from the boy's voice).

The design of filters, whether analog or digital, is a vast subject in itself. Just as in a simple electronic circuit it is far from intuitive to decide what arrangement of capacitance, inductance and resistance will provide a specified frequency response, so the design of filters can require detailed calculation straying into the complex plane of imaginary numbers, for example.

Digital filters need not be complicated in practice, however. By simply taking a running average of pairs of successive samples, for instance, one smooths any sudden change in amplitude thereby effectively applying a low pass filter. More generally, by adding or subtracting appropriately weighted input and (previous) output samples one can apply a filter of almost any desired characteristics. Finding the right arithmetic combination is the problem. An alternative approach, adopted in the CHANT program[10], is to use each pulse of the signal from the sound source to generate, in parallel, the calculated impulse responses of the desired filters. These responses, which take the form of exponentially damped reverberations decaying over several sample periods, can simply be read from memory and appropriately summed for each output sample.

Perhaps the most striking feature in *Mortuos* is the apparent ease with which, with the help of these digital techniques, two totally different sources of sound (one inflexible and metallic, the other variable and human) are bonded together as complementary acoustic entities, exhibiting or merging their individual characteristics at the composer's will.

Computation in performance

Répons, the latest composition of Pierre Boulez, is something of a landmark, not least for the degree of technical and musical integration between computer and musicians. As is typical of a Boulez composition, the piece has driven music critics to extremes: "... spasmodic gasps and lurches ... no purpose ... at odd moments it was superficially stimulating as a cold shower can be..." (*New York Times*); "... a great blinding sun of a piece, commanding its satellites in majestic orbit ... the supreme exercise of the intellect becomes the supreme gratification of the senses ... reverberating like bell towers of a mighty celebration" (*Boston Globe*). As it happens, no-one (other than the composer, perhaps) has heard the work in its entirety. The musical score is complete, however, and many of the computer transformations have been written, and have been heard in a number of preliminary performances.

The title of the piece, derived from the words *réponse* (reply) and *réponsoire* (responsory), reflects the composer's focus on relationships between a soloist and a group. This idea applies both literally to the ensemble of performers—six soloists responding to a 24-strong group of instrumentalists, one computer responding to the soloists—but also figuratively to the music itself. The aural clarity of much of the writing and the sheer drama of many of the computer transformations combine to yield a formidable impact in the concert hall.

The physical layout of a performance

of *Répons* is itself unusual. Six soloists (two pianos, harp, vibraphone, glockenspiel and cymbalom) and six loudspeakers are distributed around the performing space while the intrumental ensemble (woodwind, brass and strings), electronic processors and conductor are placed centrally. Only the soloists, connected to the processors by acoustic and contact microphones, are electronically transformed although their natural sounds are, of course, also heard.

The soloists have certain sound characteristics in common. Each note, being produced by struck or plucked strings or a metal bar, consists of a sudden attack and resonant decay, such that every instrument can be subjected to the same type of computer transformation. The 13,000 or so lines of computer code already used in the piece produce a variety of transformations such as modulation of one instrumental tone by another, controlled retardations to produce rhythmic patterns, frequency shifting and selective amplification and mobile distribution of sounds around the speaker system. (Boulez specifies very precisely the effects that he wants. His instructions are then expressed in computer code by his assistant Andrew Gerzso.)

The technology of *Répons*

The signals from the six soloists feed via a standard mixer into a 'matrix'. This 16 × 16 network can be programmed to adopt particular configurations by which, for example, one of the input channels can be fed into several outputs or vice-versa; furthermore these analog channels can each be amplified or attenuated. The matrix, controlled by a 68000 VME computer, distributes processed signals around the speaker system, again according to preprogrammed patterns which are activated as the performance proceeds. The matrix is also responsible for feeding signals to and from the analog-to-digital (AD) and digital-to-analog (DA) converters which form the interface with the core of the system, the 4X sound processor which is itself also controlled by a 68000 computer.

Designed by ex-nuclear physicist Giuseppe di Giugno, the 4X is the most powerful sound processor currently available, the speed and scope of which stem from its eight separately programmable processor boards[11]. Each of these consists of a unit for arithmetic and logical operations (the latter manipulating the 24 individual bits that make up each binary word, the former using those words as standard binary numbers), a multiplier and three main memories: 64,000 words, each of 16 bits, for the wave-table, 1,024 words of 48 bits for program storage and 1,000 words of 24 bits for program variables. Also within the 4X there is a control board, including 256 programmable timers which can, for example, initiate exchanges of program segments or memory with the 68000 host computer at prearranged times.

For all its sophistication, the 4X has a number of shortcomings which the next generation of machine—the 5A, scheduled for completion in 1989—should overcome. As well as being physically more compact and portable as a result of its use of very large scale integrated circuits, the 5A will have better data width or resolution (32 as opposed to 16 bits), a higher sampling rate (40 rather than 16 kHz) and larger storage capacity.

A more subtle improvement will follow from the use of floating point rather than integer arithmetic. Calculations involving floating point representation are slower but the improvements in signal-to-noise ratio arising from increasing resolution (or decreasing quantization error) are more substantial. The extra data width will provide another benefit: although the resolution of the AD/DA converters may be 16 bits (which is entirely adequate for most purposes) a higher resolution is needed within the calculating units to minimise the effects of rounding errors which accumulate over successive arithmetic operations, potentially disastrously if only 16 bits are used.

For any particular section of the musical score of *Répons* there is a corresponding segment of computer program, keyed in at the appropriate moment of the performance by typing in a single stroke at the keyboard of the 68000 VME. Throughout that section of the piece, while the players and conductor perform the musical score in the usual fashion, any signal from a particular soloist will be transformed automatically by the respective 4X processing board, subject to independent routing or attenuation by the matrix. Boulez intends to develop more sophisticated and multipath programs for subsequent compositions.

The music of *Répons*

Of all the computer-assisted musical gestures used by Boulez, by far the most arresting and theatrical occurs at the point where the soloists and the 4X first enter the proceedings (score number 21 in ref. 12). This entry follows an extended introductory exposition in which Boulez presents musical metaphors for the title of the piece, including most noticeably sections where certain pitches attain particular prominence while chords and textures shift about them—a procedure equivalent to the 'pedal' in classical harmony. The subsequent music sets the instrumental ensemble and the soloists in a continual dialogue, the two groups performing contrasting musical materials.

This pedal device again assumes prominence at the soloists' entry, represented by a massive chord played by all six in arpeggiated fashion (that is, the notes contributed by each soloist are struck or plucked in rapid ascending order). Different soloists are amplified through different speaker combinations adding to the spatial impact of the gesture. As this chord decays, successive soloists each play an answering chord, the respective chords having top and bottom notes in common (see Fig. 4), adding a degree of aural focus to the electronic manipulation which is also applied at this point.

Two computer transformations operate here: time delays and frequency shifts. The delays are applied simply by storing every digital sample into a wavetable as it arrives and then retarding the "pointer" by an amount corresponding to the milliseconds of delay required. Each delayed signal can be individually attenuated.

A selection of these delayed signals are fed through a variety of digital frequency shifters. By appropriate manipulation of a group of successive samples these add or subtract a specified frequency to every harmonic component of the signal[9].

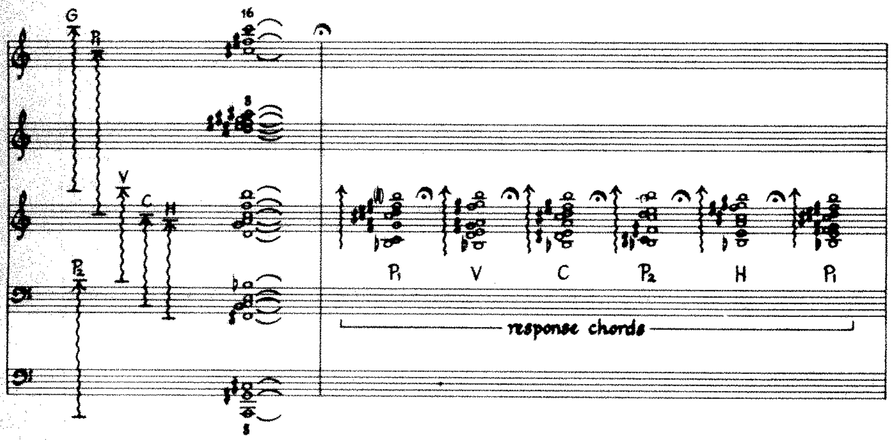


Fig. 4 Entry of soloists in *Répons*, reduced score. All chords are played as loudly as possible, each separated by several seconds. The arpeggiated contributions of each soloist to the initial chord are indicated: P1,2, pianos 1 and 2; H, harp; V, vibraphone; C, cymbalom; G, glockenspiel. Note that the part of the initial chord played on the vibraphone is of particular significance (see text). The six chords that follow are subjected to extended computer transformations.

Because the musical scale is logarithmic (for example, every octave step upwards in pitch represents a doubling of frequency), the linear transformation of frequency addition or subtraction will completely alter the harmonic relationships and hence the character or timbre of the sound. In many cases the result is reminiscent of tuned metal percussion instruments from the Far East—the Balinese *Bonang* or the Chinese *Yün-lo.* Despite the diminished "purity" of the spectrum, the fundamental pitch is usually clearly audible.

The manipulation of pitches, particularly within chords, is fundamental to the musical and electronic writing in *Répons.* One chord in particular appears to contain the roots of much which follows: that played by the vibraphone within the first entry chord (see Fig. 4). This appears in a prominent form at the very opening of the piece; moreover the surrounding components of the entry-chord, each of the answering chords played by the soloists and some of the frequency shifts are derived directly from it[13]. This 'primary' chord also appears prominently elsewhere.

An attempt to represent the combined effect of delays and frequency shifts, as specified by the *Répons* software for one of the soloist's response chords, is shown in Fig. 5.

These techniques are applied at several other points in the composition but the aural images that result vary considerably according to the musical context. For instance, at one point all the soloists are gathered together (in the bar preceding number 27 in ref. 12) to play the primary chord discussed above. This recapitulation of the chord is of some significance within the overall form of the piece and is certainly exhibited as such: it suddenly emerges at full volume from the ensemble of soloists, being sustained by tremolando with gradually decreasing loudness over several seconds—a significant duration in this context. The solo tremolandos are all fed through the same sound processor with several delays being applied, each delay with its respective frequency shift and each directed through a different speaker. Thus the scope of this already powerful moment is enhanced in pitch and space.

One technique not mentioned so far, and yet one which dominated earlier developments in analog music studios, is ring modulation. Here two signals are simply multiplied together. For two sine waves this results in the suppression of the two waves and the generation of two product waves or sidebands whose frequencies are the sum and the difference of the originals. For any instrumental sound one effectively multiplies every harmonic component by a modulating sine wave of specified frequency, which can result in very "thick" tones, packed with partials (spectral components which are not harmonically related to the fundamental) and bearing no resemblance to the original tone. The principal interest of ring modulation arises from the subtlety in the choice of the modulating frequency. If this frequency is harmonically linked to those of the instrumental tone, the resulting partials overlap to yield a 'cleaner' spectrum.

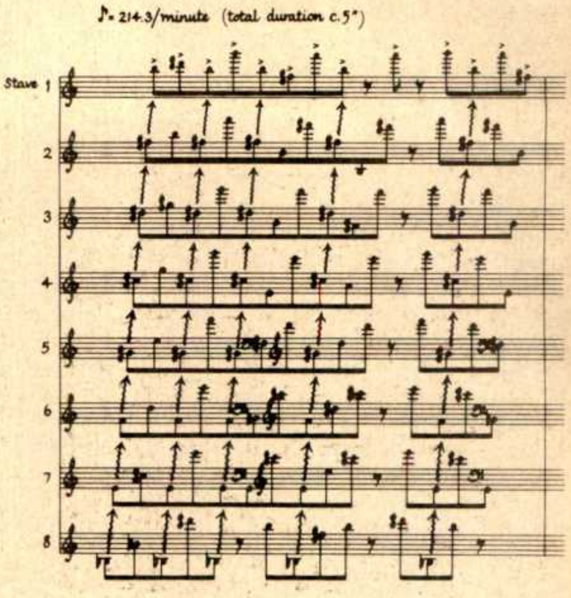


Fig. 5 Computer transformation generated following the final chord shown in Fig. 4. Each stave (from 1 to 8) contains the notes (or their nearest semitone approximations) that are generated by each of the eight notes in the chord. Thus the first note of each stave—linked from stave to stave by arrows—forms part of the arpeggiated chord played on the piano. Subsequent notes linked in similar fashion form the same chord delayed without frequency shift. All other notes are the products of delays and frequency shifts. The 14 delays applied are regularly spaced at 0.28 seconds. The uppermost initial note is emphasized by the pianist in performance, so that its transformations dominate the texture. The ear does not follow each line but resolves the texture into layers of rhythmic patterns at different registers (bass, treble and so on). This combined with the exotic timbres resulting from the frequency shift gives an overall effect of multiple echoes of the piano's chord in the midst of a rhythmic gamelan-like texture. This diagram was derived from the computer code of *Répons.*

One particularly complex example of this approach occurs in *Répons* (at score number 53, ref. 12). Here Boulez gives each soloist a different collection of seven pitches to be played in a variety of orders and rhythms. In designing the ring modulations for this passage, the composer treated these notes as eighth harmonics of respective fundamentals, choosing as modulating frequencies the *n*th harmonics where $n = 2$ for the first piano, 3 for the vibraphone and so on. The groups of modulating frequencies are injected on the respective processing boards of the 4X. Simultaneously, Boulez specifies several optional layers of transformations with differing combinations of delay and ring modulating frequencies. The options vary from instrument to instrument, the particular combination of layers being chosen in part by a random number generator, resulting in the most elaborate transformations in the entire piece.

Such complexity, which if excessively applied leads to tired ears however fascinating the detail, is preceded and contrasted by a simpler but still striking device, in which every soloist rapidly reiterates (albeit with ornamentation) a single note (B below middle C—see score number 52, ref. 12), thus providing an excellent opportunity to throw the computer's role into sharp definition. Here it acts as a ring-modulating 'light-house': a series of 'flip-flop' switches within the software scans around the soloists, applying time-varying modulation frequencies giving rise to rapidly varying spatial distributions and bands of transformed tones diverging upwards and downwards in pitch.

The final moments of this extensive piece refer back to the first entrance of the soloists discussed above, both in its musical elements and in the combination of frequency shifts and delays used by the computer. As was shown in Fig. 5, the use of arpeggio in the soloists' chordal responses gave rise to markedly rhythmic sequences of notes overlain out of phase. Here, in contrast, the transformed chords are played without arpeggio. The final gesture of all, for example, comes from the computer as it reflects—via a few delays and frequency shifts—a single quiet piano chord. The lack of arpeggio in the initial attack results in a well-defined (albeit brief) rhythmic series of transformed chords, a surprisingly simple and gentle end to a massive and complex composition, and all the more effective for that.

The past, the future

This essay may be considered by some to be something of an insult to the general discipline of computers and composition—no serious attempt has been made, after all, to set *Mortuos* and *Répons* within their historical and contemporary context. For example, the work of the Italian futurists pales into complete insignificance when compared with that of the equally radical pioneer of *musique concrete,* Edgard Varèse. He, at least, held a vision whose fulfilment may now be seen on the horizon. "I dream of instruments obedient to my thought" he wrote, "and which, with their contribution of a whole new world of unsuspected sounds will lend themselves to the exigencies of my inner rhythm". And for years such centres as Stanford, MIT, Bell Laboratories and institutes in Paris, Cologne and Milan have developed the hardware, the software and the music on which this fulfilment is being built.

For all that, it seems that such studies have reached a turning point within the general context of contemporary music. A substantial proportion of the most gifted young composers of the 1950s and 1960s were obsessed with musical language (and necessarily so), to the extent that uncompromising explorations alienated the bulk of the public audience traditionally associated with 'serious' or 'art' music. Only a few, the most notable being Karlheinz Stockhausen, found the world of electronics appropriate to their modes of communication (or, in too many cases, the lack of it). Thus the combinations of exploratory languages and techniques pushed composition with electronics into a ghetto within a ghetto.

In contrast, the powerful techniques becoming available for digital synthesis, and compositional achievements such as those of Harvey and Boulez at IRCAM, have recently stimulated a degree of interest among the wider community of composers that is unprecendented. Given the technical facilities and assistance that IRCAM is now providing to composers new to electronic techniques, and given that many of these composers are writing music that is less forbidding than much written only fifteen years ago, the time must rapidly be approaching when the widest musical public will be able to encounter the best of this work with ease.

This enhanced degree of interest from composers is likely to throw into relief all the more vividly the questions—both practical and theoretical—that are raised by the use of computers. For instance, what is their potential formal significance in live transformation? The devices used in *Répons*, for example, are purely passive in formal terms in that they produce reflections of composed material. These aural images add an indispensible component to the fabric of the piece but not to its essential form—they are the equivalent of a cathedral's windows rather than its weight-bearing walls and arches.

But the potential is there. As a move in that direction, Boulez intends to return to a previous composition, Explosante-Fix, which incorporated (unsatisfactorily in his view) live electronic manipulation. The piece will be revised so as to include reactive manipulation of the structures of chords as the performance progresses.

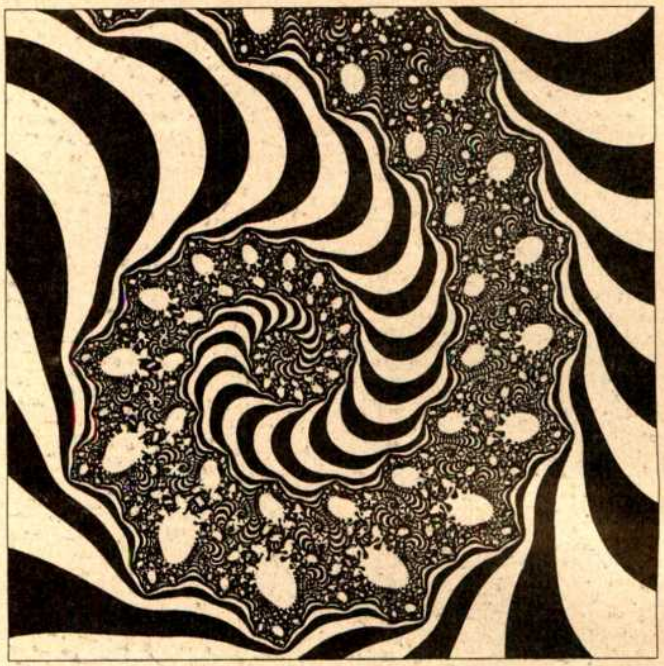


Fig. 6 A highly expanded view showing the border of the Mandelbrot set for the operation $x_{n+1} = f(x_n) = x_n^2 + c$, as mapped onto the complex c plane. The black and white areas demarcate values of c that lead the process $f(x)$ to different regions ('attractors') of the complex x plane[14].

And there are more extreme possibilities. One can imagine an abstract composition whose micro- and macro-structures derive entirely from mathematical formulae exploiting relationships between pitches, harmonics and rhythms for example. Before dismissing such a scheme as depressingly arid, the sceptical reader might examine pictorial representations derived from a mathematical entity known as the Mandelbrot set[14]. The nonlinear properties of the associated formulae, mapped onto the complex plane, produce an endless and stunning variety of spectacular images (for example see Fig. 6). A literal transcription from the two-dimensional pictorial frame to that defined by frequency and time would prove unsatisfactory. Nevertheless if there is an analogous "fractal" music waiting to be discovered then only digital techniques are likely to find it.

Other more fundamental questions have been touched upon recently by one of France's most experienced and influential computer-composers, Jean-Claude Risset[15]. For example, the access provided by digital systems to what one might call a multidimensional sonic continuum leads to a problem of aural hierarchy: to what extent does musical cognition depend on discrete structures such as scales or rhythms? And once such basic elements of language have been defined, further hierarchical problems arise, as many composers have realised following the musical revolutions that occurred in the early years of this century. Specifically, given completely new possibilities for the synthesis of sound, what are the self-consistent equivalents of the phrase, the counterpoint, the development section or the symphony that sprang naturally from diatonic and chromatic harmony in earlier times?

No doubt the strongest of musical instincts will continue to be required to solve such problems without sacrificing the degree of communication necessary to justify the effort. "It has often been said that music is as much a science as an art;" writes Boulez[16]. "How could these two entities be fired in the same crucible, except by the imagination, that queen of faculties!".

Thanks are due to Pierre Boulez for access to IRCAM, to Andrew Gerzso for introducing me to the workings of *Répons* and for supplying scores and software, to Gerzso and Jonathan Harvey for helpful comments on the article, and to Elaine Gould for her musical calligraphy. ☐

Philip Campbell is Physical Sciences Editor of Nature.

1. Xenakis, I. *Formalized Music* (Indiana University Press, Bloomington, 1971).
2. Longuet-Higgins, H. C. *Nature* **263**, 646–653 (1976).
3. Erato, STU 71544.
4. Harvey, J. *Computer Music J.* **5**, 22–24 (1981).
5. Dodge, C. & Jerse, T. A. *Computer Music* (Schirmer, New York, 1985).
6. Moorer, J. *Computer Music J.* **1**, 4–37 (1977).
7. De Poli, G. *Computer Music J.* **7**, 8–26 (1983).
8. Mathews, M. V. *The Technology of Computer Music* (MIT Press, Cambridge, Massachusetts 1969).
9. Potacsek, P. & Favreau, E. *Manuel d'Utilisation du Language de Patch* (IRCAM, Paris, 1985).
10. Rodet, X. Potard, Y. & Barriere, J.-B. *Computer Music J.* **8**, 15–31 (1984).
11. Koechlin, O. *et al. IRCAM Rapport Recherche*, 39 (1985).
12. Boulez, P. *Répons* (Universal Edition, 1986).
13. Gerzso, A. *Contemporary Music Rev.* I, 23–34 (1984).
14. Peitgen, H.-O. & Richter, P. H., *The Beauty of Fractals*, (Springer, Berlin, 1986).
15. Risset, J.-C. in *Composers and the Computer* (ed. Roads, C.) (Kaufman, California 1985).
16. Boulez, P. in *Boulez on Music Today*, 143 (Faber and Faber, London, 1971).

Cosmic microwave background anisotropy

Nick Kaiser* **& Joseph Silk**†

* Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK
† Astronomy Department, University of California, Berkeley, California 94720, USA

Current hypotheses for the origin of structure in the Universe lead to predictions of the amplitudes of anisotropies in the cosmic microwave background radiation. The dipole anisotropy is related to density fluctuations on large scales and to other determinations of our motion relative to distant galaxies. Observation and theory are coming tantalizingly close to measuring the elusive anisotropy, or to revealing that our ideas about the origin of galaxies and large-scale structures are in need of substantial revision.

STUDIES of the angular structure of the microwave background radiation play two distinct roles in cosmology. First, there are those anisotropies that are intrinsic to the microwave background. These provide an unsurpassed glimpse of the early Universe, propagating freely to us from an epoch when the Universe was much simpler than it is today. Galaxy clusters had not formed, nor, most probably, had galaxies themselves. Detection of anisotropies in the background radiation would provide a unique measure of the primordial density fluctuations from which the present large-scale structure of the Universe developed. Yet despite continuing and ever-improving measurements, no such anisotropies have been discovered. Here we describe some of the current hypotheses for the origin of structure in the Universe, and the calculations which have been made to predict the amplitude of the microwave anisotropy. A different test of cosmological theories is provided by the dipole anisotropy of the microwave background, which arises from our motion relative to a uniformly expanding 'cosmic frame' defined by the mean state of motion of matter at great distances. We describe this anisotropy as extrinsic, and infer a solar motion which can be used to constrain both the large-scale density fluctuations in which we are embedded and independent determinations of our motion relative to distant galaxies. In the latter part of this review we describe the present status of these studies.

Origin of structure

Large-scale structure originated, according to the primaeval hypothesis, from infinitesimal density fluctuations in the earliest instants of the Big Bang. Inflation[1] has provided a tentative answer to the question of the origin of these fluctuations[2-5]. Quantum fluctuations are amplified to large scales during the exponential expansion phase which characterizes the inflationary epoch. The inflation is driven by vacuum energy during a first-order phase transition which occurs as the energy of matter decays. The universe evolves from the symmetric state of false vacuum, when the fundamental forces are unified, to the asymmetric state of the true vacuum. During the inflationary phase, described by de Sitter space-time, there is no particle horizon. The fluctuations are boosted to a scale-invariant amplitude, up to scales comparable to the radius of the inflated space-time bubble, which greatly exceeds our present horizon, and models can be constructed that yield density fluctuations at horizon crossing of

$$(\delta\rho/\rho)_{\mathrm{H}} \approx 10^{-4} \qquad (1)$$

Values much larger ($(\delta\rho/\rho)_{\mathrm{H}} \approx 1$) would be disastrous for cosmology, producing black holes and large inhomogeneities, and much smaller values would not lead to any galaxy formation. It is remarkable that constraints on the precise value of $(\delta\rho/\rho)$ can be inferred from constraints on temperature fluctuations $\delta T/T$ in the cosmic microwave background. Indeed, all currently acceptable cosmological theories require large-scale structure to have originated from infinitesimal fluctuations at early phases of the Big Bang, and these inevitably leave a signature in the form of temperature fluctuations $\delta T/T$. Whether the amplitude (1) will emerge naturally from a quantum gravity model is not known. Because quantum gravity is poorly understood, most approaches to this problem begin when a classical description first becomes valid, well below the Planck energy of $\sim 10^{19}$ GeV. According to one school of thought, the inflationary potential can be designed to yield the correct $\delta\rho/\rho$, which can then be interpreted as an otherwise unknown particle physics parameter. This process results in a scale-invariant spectrum of gaussian density fluctuations. Another view is that the initial fields that describe the energy content of the Universe are chaotic and highly variable in space and time[6]. Beginning with this initial space-time foam, an anthropic argument is applied to infer that a very special region, with sufficiently regular fields, will inevitably exist and be able to inflate sufficiently to dominate space-time, producing a highly isotropic (and nearly homogeneous) universe. More than one epoch of inflation could have occurred[7], and a wide range of possible primordial fluctuation spectra can arise.

One intriguing consequence is that the primordial fluctuations might be formed exclusively on small scales, far smaller than those of galaxies or galaxy clusters. Galaxy formation can still occur, however, if nonlinear structures develop on stellar mass scales; an amplification process can then produce large-scale structure[8,9]. This happens as follows: massive stars form, explode as supernovae, and sweep up large shells of ambient matter which subsequently fragment on much larger scales, into pieces with the mass of star clusters. If sufficient energy is to be released by dying stars at this stage for the process to repeat, amplifying to larger and larger scales, then formation of structure on galactic mass scales can be achieved. Galactic binding energy, in such schemes, may be said to be of mediaeval, rather than primaeval, origin: it is secondary, having been induced by the explosions altering the distribution of matter.

In the primaeval picture, the excess binding energy of a cluster of galaxies can be traced back to a primordial excess of density in the same co-moving region at the inflationary epoch, whereas in the mediaeval picture this binding energy is imposed at a very late epoch, when the region is well inside the horizon. There is yet another hypothesis, the 'cosmic string' model[10,11], in which the binding energy fluctuations on a given scale are induced by the dynamics of the string network as that scale enters the horizon.

In each of these models there is a link between the very earliest instants of the Big Bang and the large-scale structure of the Universe today. The microwave background fluctuations result from the linear growth regime of the primordial fluctuations, and thus bypass much of the uncertainty associated with nonlinear evolution. Observations of the microwave background anisotropy provide a unique probe which may help to distinguish between the various models for the origin of galaxies.

Intrinsic anisotropy

Figure 1 shows what one may learn by studying the anisotropy of the cosmic microwave radiation. The microwave photons stream freely to the observer ('here and now') from the surface of last scattering, which in the standard model of the Big Bang occurs in the redshift range $z \approx 1{,}100$–$1{,}500$. The last scattering surface has a finite thickness Δz, which is due to the noninstantaneous recombination of the hydrogen and to the residual ionization level which freezes out. A transmission factor per unit redshift interval, $(d\tau/dz)\exp(-\tau)$, defines the effective profile of the last scattering surface, where $\tau = \int_0^z n_e \sigma_T \, c dt$ is the optical depth to redshift z, σ_T is the Thomson cross-section and n_e is the electron density. This function can be approximated by a gausian which peaks at $z = 1{,}065$, with a gaussian width $\Delta z = 80$ (ref. 12). The corresponding co-moving distance that describes the thickness of the last scattering surface is $\Delta L \approx 7h^{-1}\,\Omega^{-1/2}$ Mpc (where h is the Hubble constant divided by 100 km s^{-1} Mpc^{-1}).

The finite width of this shell smooths out fluctuations on angular scales $\Delta\theta \lesssim 8\,\Omega^{-1/2}$ arc min. Even larger-scale primordial fluctuations are not necessarily preserved: a source of energetic photons or heat could re-ionize the Universe or increase its ionization over the low post-recombination value ($n_e/n_H \lesssim 10^{-3}$) and thereby shift the last scattering surface towards lower redshift. As the thermal history of the Universe at the relevant epoch ($1{,}000 \gtrsim z \gtrsim 30$) is uncertain, there is considerable ambiguity in the interpretation of fine-scale anisotropy. This ambiguity is avoided if we search over angular scales exceeding that subtended by the particle horizon at the latest possible epoch of last scattering. No causal process could erase such structure. This angular scale is $\sim(\Omega/z)^{1/2}$ radians, where Ω is the cosmological density parameter. As the minimum redshift at which significant scattering could have occurred ($\tau \gtrsim 1$) is ~ 30 if the baryon density parameter $\Omega_b \lesssim 0.1$, we infer that an observational strategy aimed at fluctuations over angular scales of $\gtrsim 10°$ would involve the least number of assumptions about the early evolution of the Universe.

The corresponding co-moving linear scale is ~1,000 Mpc. At recombination, a linear scale of 100 Mpc subtends an angle of $\sim\Omega h$ degrees. Observations of the background radiation fluctuations therefore nicely complement studies of galaxy clustering, which provide firm evidence for structure only on scales of up to a few Mpc, in probing the large-scale structure of the Universe. Thus, the interpretation of large-angle anisotropy is independent of the uncertain ionization history, but is crucially dependent on the assumed extrapolation of density fluctuations to large scales by means of equation (1). For small-scale anisotropy the situation is reversed. Regardless of the angular scale considered, the radiation has streamed freely since the Universe was very homogeneous, at which time it was well described by linear fluctuation theory. Thus, in contrast with studies of galaxy formation and clustering[13-16], the complexities of nonlinear evolution are completely bypassed. The detection of anisotropy on any angular scale could provide a unique and direct measure of the primordial fluctuations from which large-scale structure evolved.

Predicting $\Delta T/T$. A variety of calculations predicting $\Delta T/T$ can be found in the recent literature; they differ mainly in the assumed material content of the Universe, the assumed initial fluctuations and the normalization to present-day structure. In order to make sense of the issues here, we first describe one set of calculations, which we refer to somewhat arbitrarily as the standard model, and then describe variations on this theme.

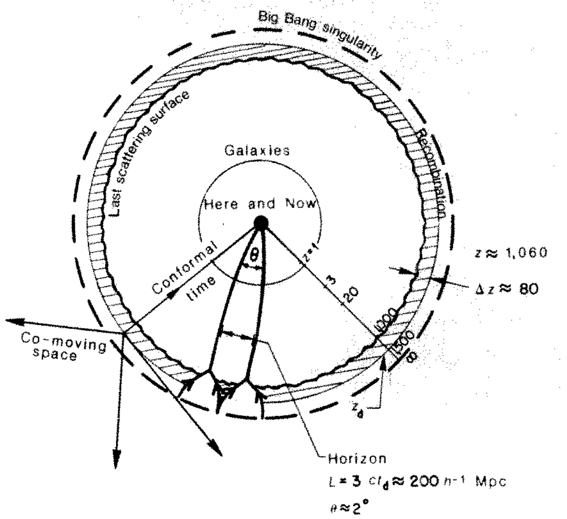


Fig. 1 A co-moving space/conformal time diagram of the Big Bang. The observer ('here and now') is at the centre; the Big Bang singularity is the outermost dashed circle, and the horizon scale is schematically indicated at last scattering.

The 'standard' model: cold dark matter with isentropic gaussian fluctuations. The fluctuations in this model have been studied in detail[17,18]. The Universe is assumed to have three components: a cold, collisionless fluid (this might represent axions, for example); the baryonic gas, initially hot and fully ionized; and the photons and neutrinos, described by a distribution function, with the former being coupled to the gas by electron scattering. The initial state is assumed to be isentropic, which means that, for any two types of particle A and B, the ratio of the number density of As to Bs, n_A/n_B, is a constant independent of position. This terminology arose in the context of a universe containing only radiation and plasma, for which the ratio of the number of photons to the number of baryons n_γ/n_B is a measure of the entropy (per baryon). Such perturbations are sometimes referred to as 'adiabatic' because they can be generated by compressing all of the particles in some region, producing a net density perturbation $\delta\rho(r) \equiv \rho(r) - \rho$. Strictly speaking, one must also specify the initial peculiar velocity field, but this is usually dispensed with by stipulating that the perturbation should contain only growing modes.

A 'gaussian' field is one which can be generated as a linear sum of spatial Fourier components with random phases. Such a field is specified by its power spectrum, which in this case is taken to be the Zel'dovich spectrum, $|\delta_k|^2 = |\delta\rho/\rho|_k^2 = Ak$ (where k is the wavenumber of a Fourier component and A is a constant to be specified by a normalization described below), which results in binding energy fluctuations that are scale-invariant.

With the initial conditions thus specified, the spatial Fourier modes can evolve independently until some time after recombination. This results in final linear fluctuations in the matter and radiation which are now decoupled. The angular dependence of the sky fluctuations is now completely specified. The amplitude, however, remains to be fixed. This 'normalization' is performed by requiring that the density fluctuations be of sufficient amplitude to generate the observed large-scale structure. This is conventionally measured by the integral of the galaxy–galaxy correlation function: $J_3(x) = 3/x^3 \int_0^x \xi_{\text{gal-gal}}(r) r^2\, dr$, where x is a measure of the separation between an arbitrarily selected pair of galaxies. Note that despite the fact that $\xi(r) \gg 1$ on small scales, this integral satisfies linear theory[19]. One evaluates this at a scale where the fluctuations are linear but large enough so that $\xi(r)$ is well known. For example, at $x = 10h^{-1}$ Mpc, the r.m.s. fluctuations are ~50%. This normalization assumes that the galaxies are 'fair tracers' of the density field. Although this is a fairly natural assumption, this need not necessarily be the case. If the Universe has closure density ($\Omega = 1$), then it seems that galaxies must be more strongly clustered than mass, at least in the dense clusters where virial analysis is applied to estimate mass to light ratios. In considering these models, some allowance must be made for this, somewhat reducing the amplitude of $\delta T/T$.

There are two final complications. Uncertainty in the Hubble constant H_0 leads to uncertainty in $\delta T/T$. Specifically, for cold dark matter, a decrease in H_0 leads to an increase in L_{eq} (the radius of the co-moving sphere encompassing the causal horizon at the epoch of equality of matter and radiation densities), and thus to an increase in $\delta T/T$. Similarly, if $\Omega < 1$, $\delta T/T$ is increased, approximately in proportion to Ω^{-1}. This is because the Universe becomes curvature-dominated at redshift $\sim 1/\Omega$, and fluctuation growth effectively ceases. One is now in a position where $\delta T/T$ can be predicted. The detailed coupling of matter and radiation must, of course, be followed. This involves solving the perturbed Boltzmann equation for the photon intensity today[20,21]. Rather than give any details here, we will simply indicate in a qualitative manner the magnitude of the various terms that contribute to $\delta T/T$.

On small scales, corresponding to a co-moving scale that subtends the thickness of the last scattering surface at ~ 10 Mpc, there are adiabatic temperature fluctuations. These were first predicted in 1967[22,23], and would be of order $(1/3)\delta\rho/\rho$ were recombination instantaneous and complete. Their angular scale is typically 5–10 arc min, and fluctuations of smaller angular scale do not survive recombination. The amplitude of these small-scale fluctuations is reduced, because of the finite thickness of the last scattering (LS) surface, to $\delta T/T \approx (\delta\rho/\rho)_{LS}/30$. On larger scales, the random motions of fluctuations induce Doppler shifts. These dominate $\delta T/T$ at co-moving scales $d \approx L_{rec}$ (the radius of the co-moving sphere encompassing the causal horizon at recombination) at $\sim$50–100 Mpc, corresponding to angular scales of 1–2°, and are of order $\delta T/T \approx (\delta\rho/\rho)_{LS}(d/ct)_{LS}$, where c is the speed of light and t is the time elapsed since the Big Bang. On the largest scales, the gravitational potential difference between last scattering and now generates gravitational redshifts, of order $\delta T/T = (\frac{1}{3})(\delta\rho/\rho)_0(d/ct)_0^2$. These fluctuations, also first predicted in 1967[24], are contributed by co-moving scales d of 10^3–10^4 Mpc, and dominate over angular scales greater than a few degrees[18].

The result of all this is a gaussian sky pattern (the gaussian nature of the fluctuations being preserved in linear theory), with roughly equal power in each logarithmic interval of angular wavenumber. The r.m.s. amplitude is a function of global cosmological parameters such as Ω, h and Ω_b, and constraints can be placed on these by comparison with the stringent small-scale anisotropy limits. High-density models are generally consistent with these limits, but low-density models can be excluded. There are two reasons for this: first, there is reduced growth of $\Delta\rho/\rho$ at late times if $\Omega \ll 1$, and second, the fluctuations appear at a more favourable angular scale. The results we have described were obtained numerically. The anisotropy on large scales, much greater than the horizon size at decoupling, for these isentropic fluctuations can be calculated analytically[24] (see ref. 25 for an application to cold dark matter). The effect is analogous to the usual gravitational redshift effect, with photons being redshifted (blueshifted) if they were last scattered in overdense (underdense) regions. The amplitude of $\Delta T/T$ is one third of the newtonian peculiar potential ϕ evaluated at the point of emission.

In these calculations it is assumed that the Universe is not re-ionized after recombination. There is some justification for this in that, for most reasonable choices of parameters, the spectrum of density fluctuations is such that very little matter forms nonlinear structure at a sufficiently early epoch to plausibly re-ionize the Universe[17,18]. We now describe some variations on this 'standard' model.

Hot dark matter. An interesting alternative is to replace the cold dark matter (axions, photinos, or the like) with hot dark matter such as massive neutrinos. The main effect of this is to erase post-recombination density fluctuations on scales smaller than that of superclusters (the angular dependence of $\Delta T/T$ being largely unaffected because the $\Delta T/T$ fluctuations on small scales are in any case smeared out)[17,26]. The normalization to J_3 (for galaxies) is now questionable, because galaxy formation is a secondary phenomenon in a neutrino-dominated universe, arising from pancake fragmentation[27]. An alternative normalization is to require that 'pancake' collapse should occur sufficiently early to form quasars and the most distant galaxies. For a neutrino mass of $\sim$30 eV, the predicted anisotropy is compatible with the current upper limits.

Isocurvature fluctuations. An alternative to the isentropic fluctuations discussed so far is the 'isocurvature' mode. These fluctuations are, in a sense, orthogonal to the isentropic modes because, while the latter have fluctuations in the total energy (and therefore in space curvature) but no variation in the relative abundances of the various particles, fluctuations can be realized by spatially varying the equation of state on some initial $t =$ constant hypersurface in a universe which was hitherto absolutely homogeneous. This process conserves energy locally, so any excess in one species is balanced by a deficit in the others. It is the constancy of the spatial curvature on this initial hypersurface which gives rise to the terminology employed here. These fluctuations are not, strictly speaking, isothermal, but are often referred to as 'entropy' or 'isothermal' fluctuations.

Isocurvature fluctuations in a universe dominated by cold dark matter have recently been investigated[28,29]. The small-scale anisotropy is very similar to that arising from isentropic fluctuations. The large-scale anisotropy, however, is about an order of magnitude larger. This increase is a consequence of two factors: the first is that the spectrum of density fluctuations is flatter than in the isentropic case, because no sub-horizon growth occurs in the radiation era. Thus, with the same normalization, the value of $\Delta\rho/\rho$ at large scales is about twice as large.

The origin of the other factor can be understood as follows. Imagine that, at some early time $z \gg z_{eq}$, we impose a positive perturbation to the axion density $\Delta\rho_A/\rho_A = \Delta n_A/n_A = \varepsilon$ in some large spherical region destined to enter the horizon at some time after z_{EQ}. We must also impose a smaller compensating negative perturbation in the radiation density, $\Delta\rho_\gamma/\rho_\gamma = -(\rho_A/\rho_\gamma)\varepsilon$, so that ρ_{TOTAL} is unperturbed. When this region enters the horizon, the density of the radiation will be negligible, having been redshifted, so we will have a positive perturbation to the proper mass enclosed: $(\Delta M/M)_{proper} = \Delta n_A/n_A = \varepsilon$. However, from considerations of causality, we know that the gravitational mass as registered by an external observer is unperturbed. There is no paradox here; the perturbed region must now be sitting in a potential well of depth $\phi = \varepsilon$ so that $M_{grav} = M_{proper}(1-\phi)$ is unchanged. If the perturbation were isentropic then we would see the Sachs–Wolfe term, $(\Delta T/T)_{SW} = -\phi/3 = -\varepsilon/3$. However, this perturbation is not isentropic: we initially increased the number of axions but the number of photons was almost unaltered. Thus, relative to an isentropic state, we have a negative entropy perturbation, $\Delta n_\gamma/n_\gamma = -\varepsilon$, which gives rise to a temperature fluctuation $(\Delta T/T)_{entropy} = 4(\Delta n_\gamma/n_\gamma)/3 = 4\varepsilon/3$. We see then that the usual Sachs–Wolfe term is augmented by an entropy perturbation of the same sign but four times larger, and the net result is to increase $(\Delta T/T)_{TOTAL}$ by a factor of $\sim$5. With the conventional normalization, the predicted anisotropy is incompatible with the limits of Fixen *et al.*[30].

The re-ionized Universe. If re-heating were to occur, the Universe would become transparent over approximately one expansion time at the mean redshift of last scattering. The photons we see would then arrive from a broad range of redshifts, $\Delta z \approx z_{ls}$. The temperature anistropy due to the large-scale structure in such a re-ionized universe is smaller than that predicted to arise in a un-ionized universe. This is because, along any line of sight through the cosmic photosphere, we will see fluctuations due to many regions adding incoherently and this results in statistical cancellation.

The dominant source of anisotropy in a re-ionized universe is the peculiar motion of the last scatterers. The calculation of the anisotropy is greatly simplified by two approximations which are expected to hold to high accuracy. First, at the epoch of last scattering, the density fluctuations that give rise to the currently observed large-scale structure were still in the linear regime, and second, the dynamical effect of the radiation on the matter was almost certainly negligible. Hence we need only calculate the effect of the matter on the radiation.

The statistical cancellation due to the projection of density fluctuations along the line of sight depends on the statistical nature of the density perturbation field $\delta(k)$. If this density field was gaussian (that is, if the Fourier components $\delta(k)$ had random phases), then the temperature fluctuations on the sky, $\Delta(\theta)$, will also be a gaussian field. The power spectrum of this field, Δ^2, is related to the power spectum of density fluctuations by $\Delta^2 \propto k^{-4}\delta^2(k)$, where the constant of proportionality depends on the ionization history of the Universe[31]. A consequence of the fact that the scale of the fluctuations is much smaller than the thickness of the photosphere is that the fluctuations $\Delta(k)$ at some angular wavenumber k are determined solely by the value of δ^2 at the corresponding spatial wavenumber.

A useful measure of the amplitude of the temperature fluctuations at angle θ is $(\delta T/T)_\theta \approx \sqrt{k^2\Delta^2(k)}$, where $k \approx 1/\theta$. Similarly, the amplitude of the density fluctuation on scale λ is $(\Delta\rho/\rho)_\lambda \approx [k^3\delta^2(k)]^{1/2}$. Thus, to order of magnitude, we have $(\Delta T/T) \approx (H\lambda)^{5/2}(\Delta\rho/\rho)$ (where H, λ and $\delta\rho/\rho$ are evaluated at z_{ls}). Equivalently, we can write $(\Delta T/T) \approx (1/\sqrt{N})\phi_\lambda$, where ϕ_λ is the gravitational potential fluctuation on scale λ and N is the number of regions of size $\sim\lambda$ seen along a line of sight. One can infer that the anisotropy due to fluctuations of a given scale λ is smaller for smaller z_{ls} (as N is then increased), and ϕ_λ is approximately equal to the square of the velocity dispersion of the final nonlinear structures produced. As this velocity dispersion appears to be a non-decreasing function of mass (and therefore of λ), the result tells us that the greatest anisotropy is generated by the largest structures we see today. If galaxies fairly trace the mass distribution, then the anticorrelation seen in the redshift survey suggests that the power spectrum falls rapidly for $k \leq 2\pi/60h^{-1}$ Mpc. These fluctuations would generate $\Delta T/T \approx (3\text{-}6)\times 10^{-6}$ on an angular scale of $\sim 1°$. (We have here assumed $\Omega \approx 0.2$ in accord with the assumption that galaxies trace the mass). In a high-density universe ($\Omega \approx 1$), the amplitude would be smaller by a factor of ~ 3 and the anisotropy would appear at an angular scale of ~ 10 arc min.

On larger scales, the amplitude of $\Delta T/T$ depends on the extrapolation of the power spectrum. If this is assumed to be a power law with $\delta_k^2 \propto k^n$, then $\Delta T/T \propto \theta^{(-n+2)/2}$ and would increase with angular scale if $n \leq 2$. The estimate we have given above is the minimal anisotropy implied by the structure observed today, and it does not require any extrapolation of the power spectrum to large scales. The predicted anisotropy is well below the current upper limits on the relevant scales.

So far we have not specified the source of energy which re-ionizes the Universe. An interesting possibility is that these pre-galactic sources of radiation may, if they are sufficiently inhomogeneous, generate the large-scale structure[32]. A constraint on such theories is that, in addition to the 'Doppler scattering' fluctuations we have discussed above, the radiation released would probably have been degraded to microwave frequencies and would appear anisotropic on a scale of a few degrees, this being the angular size subtended by the width of the photosphere.

This effect was investigated in ref. 33. The assumptions are that a population of sources turn on and burn at redshift z_*, that these sources are inhomogeneous on some scale λ_* and are uncorrelated on larger scales (note that there is no assumed initial density inhomogeneity), and that these sources produce a substantial fraction of the microwave background. The diffusion or streaming of this radiation away from the sources then generates density fluctuations by the pressure of the radiation of the gas. This process can transfer momentum, and generate density fluctuations, on scales up to the radiation Jeans length, which is comparable to the scale of the largest structures we see today. The redshift of burning was taken[33] to be $z_* \approx$ 200-500; if z_* is lower than this, then radiation drag becomes inefficient, and for higher z the luminosity of the sources must exceed the Eddington luminosity. This process could generate the observed large-scale structure if the initial scale of radiation inhomogeneity is of the order of a galactic mass scale. The transfer of radiation smooths out the initially inhomogeneous radiation on scales up to the horizon scale at last scattering. The final microwave background anistropy predicted (assuming thermalization of the radiation by dust) appears at an angular scale of a few degrees and is comparable to the present upper limits on this scale.

On larger scales the temperature anisotropy decreases as $\Delta T/T \propto \phi^{-1}$, and so any quadrupole anisotropy would be very small. This $\Delta T/T \propto \phi^{-1}$ behaviour at large angles is the hallmark of spontaneous theories of formation of structure. Most 'primordial' theories of large-scale structure invoke a Zel'dovich spectrum which would generate large-angle $\Delta T/T$ fluctuations which are scale-independent. Thus observation of large and comparable quadrupole, octopole and higher-order moments would rule out spontaneous generation theories. Unfortunately, our Galaxy gives rise to anisotropy, with significant low-order multipoles. It may then be that one would be able to detect the Zel'dovich anistropy only on intermediate angular scales, where one could test for a dependence of $\Delta T/T$ on galactic latitude.

Anisotropy from cosmic strings. All of the mechanisms we have discussed so far give rise to gaussian fluctuations, either by virtue of the assumed initial conditions or, as in the re-ionized universe, where we have many fluctuation regions adding incoherently, so that the central limit theorem should apply. An interesting alternative is supplied by the 'cosmic string' hypothesis[10,11] (for a review of this subject see ref. 34). The perturbations provided by cosmic strings are inherently non-gaussian and result in step-like discontinuities snaking across the microwave sky[35-37]. The pattern is like that arising from a Zel'dovich spectrum, scale-invariant on large scales. In this picture, one would naturally expect the Universe to be re-ionized, so the theory is relatively immune to the constraints from small-scale anisotropy. In order to detect the distinctive features predicted in this model, one would need to scan a region of sky at least $\sim 10°$ square with reasonably good resolution.

ΔT/T *in the explosive model.* Explosive amplification[8,9] begins with rare seeds, either massive objects ($\sim 10^6$ solar masses) or possibly galaxies themselves, and results in a foam-like distribution of galaxies: swept-out voids separate thin shells, sheets and ridges where shells intersect, and there most of the galaxies form. The typical void size can be up to ~ 5 Mpc in diameter if sufficient explosive energy is available. Although this falls short of the observed voids in the galaxy distribution, which are ~ 25 Mpc across[38], the qualitative resemblance of the bubble-like structure of the galaxy distribution to that predicted in explosive amplification is tantalizing, and this model is a serious contender for explaining galaxy formation.

In fact, the voids are not completely empty, but contain hot shocked gas at a temperature of $\geq 10^8$ K. In the pancake fragmentation model, the physics of galaxy formation is very similar to that in the Ostriker-Cowie-Ikeuchi model[8,9]. Pancakes collapse to form thin, dense sheets of gas which fragment into galaxies, leaving behind large regions of tenuous hot gas. This hot gas can result in observable fluctuations in the cosmic microwave background, by means of the Sunyaev-Zel'dovich effect[39]. Microwave photons Compton-scatter against the hot electrons and acquire a perturbation in temperature, in the Rayleigh-Jeans part of the spectrum, of $\Delta T/T \approx -(2kT/m_e c^2)\tau_{\mathrm{es}}$, where k is the Boltzmann constant, m_e is the electron mass and τ_{es} is the scattering optical depth. $\Delta T/T$ for an individual pancake is proportional to the pressure, which is constant in the shocked region and depends only on the co-moving wavelength of the pancake and the epoch of pancake collapse. Although the distortion from an individual pancake at $z \approx 1$ is small ($\Delta T/T \approx 10^{-7}$), the contribution from the many ($N > 100$) pancakes along a typical line of sight increases the variance of the temperature fluctuations by a factor of $N^{1/2}$ between different lines of sight. Hogan[40] performed a similar calculation for the explosive galaxy formation model. He found that if the observed large-scale structure as measured by the galaxy autocorrelation function is generated explosively, small-scale anisotropy comparable to the Uson-Wilkinson limit[41] is produced. Vishniac and Ostriker[42] generalized this argument further, by noting that if the binding energy of the gaseous components of newly bound structures is thermalized, regardless of its origin (whether explosive or gravitational), then temperature fluctuations of amplitude $\sim 10^{-5}$ result on sub-arc-minute scales. Primordial fluctuations are greatly diminished on these scales due to the smearing associated with the finite thickness of the last scattering surface, and such secondary fluctuations are expected to dominate.

Searches for intrinsic anisotropy. Since the discovery of the microwave background radiation[43], many attempts have been made to search for angular anisotropy. These searches have used a variety of techniques to probe a range of angular scales spanning more than three orders of magnitude. It is often convenient, in the discussion of temperature anisotropies, to decompose the sky temperature pattern into spherical harmonics; these provide a complete and orthogonal set of modes. As we have mentioned, the $l = 1$ mode is probably dominated by our locally generated peculiar velocity, which may also generate a small $l = 2$ component. Any other anisotropies must be intrinsic. Observations at different angular scales suffer from different problems and require different observing strategies; it is useful to distinguish between 'large' angular scales for which $l \leq 10$, 'intermediate' angular scales with $l \approx 10$–100, and 'small' angular scales with $l \geq 100$.

For large and intermediate angular scales it is necessary to raise the observing apparatus above the bulk of the atmosphere. Anisotropometers have been deployed from balloons[30,44–46], from U-2 aircraft[47] and from a satellite[48]. These observations have typically employed beam-widths of a few to ten degrees and have obtained coverage of a significant fraction of the whole sky. Initially there was much emphasis on the possibility of a quadrupole anisotropy, and observers obtained estimates of the quadrupole tensor components from their data. These observations also contain useful information about higher-angular-frequency fluctuations on scales down to the beam-width. Although there have been claims of positive detections of quadrupole anisotropy at levels significantly above that of the noise, none of these have withstood the test of time, and the current upper limit from the Soviet RELIC experiment on any $l = 2$ mode is $\Delta T/T \leq 3 \times 10^{-5}$ with 95% confidence. The limit on the $l = 4$ mode is 5×10^{-5} (95% confidence). These limits are much smaller than the dipole amplitude, and this fact gives some support to the belief that most, if not all, of the dipole anisotropy is extrinsic or local, since one would expect any intrinsic dipole anisotropy to be of similar amplitude to the other low-order moments.

On intermediate angular scales there has also been one claimed positive detection of anisotropy[44], although few details of the method of analysis were given and no confidence level was quoted. Other workers have quoted upper limits of $\Delta T/T \leq 3 \times 10^{-5}$. More recently, an upper limit[49] of $\Delta T/T \leq 5 \times 10^{-4}$ (95% confidence) has been set at a beam-width of 2°. An upper limit of 5×10^{-5} (95% confidence) has been set over ~6–12° in a recent dedicated experiment at 3 cm wavelength (R. D. Davies *et al.*, in preparation). At small angular scales the emission from the atmosphere is less of a problem, and searches for anisotropy on scales of a few arc min have been made using conventional ground-based radio telescopes. The comparison of the results from these various observations, and their relation to quantities which can easily be predicted from theoretical models, are hampered by the variety of observing strategies employed. In order to achieve the high sensitivity required, it is necessary to perform a differential measurement; many different choices of beam geometry have been used. One possibility would be to allow the beam to drift across the sky and measure the r.m.s. fluctuations in the receiver output. This would then tell us the r.m.s. fluctuation in the sky temperature after this has been convolved with the beam pattern. In practice, however, a more common strategy is to repeatedly switch between two or more patches of sky and subtract the signal. If only two patches are observed, the output is simply the autocorrelation function of the (convolved) sky brightness. If, as is usual, the beam throw is comparable to the beam-width, then the output is a useful measure of the fluctuations on this angular scale. Such an apparatus would also have some sensitivity to fluctuations on much larger angular scales, because the apparatus measures the gradient of long-wavelength fluctuations. The most stringent current upper limits to small-scale anisotropy are those obtained by Uson and Wilkinson[41], using a double beam subtraction technique. They quote an upper limit $\Delta T/T \leq 2 \times 10^{-5}$ at $\theta \approx 4.5$ arc min, again with 95% confidence. This type of measurement is even less sensitive to large angular fluctuations because the beam geometry measures the second derivative of the sky pattern at large angles. Finally, we mention some searches for anisotropy on 'very small' angular scales (a few arc s), using the Very Large Array[50]. The sensitivity is somewhat poorer than that obtained on larger angular scales, with upper limits of 10^{-3}, 8×10^{-4} and 5×10^{-4} (95% confidence) at angular scales of 18, 30 and 60 arc s, respectively. These results have now been superseded by new, deeper survey at 6 cm, which improves on these limits by factors of 4 and 8 at 18 and 60 arc s, respectively[51]. Similar results have been obtained independently by H. M. Martin and R. B. Partridge (in preparation).

We also expect the microwave sky to display some linear polarization. To date, polarimetry searches have been carried out only at large angular scales[52,53], these searches having been stimulated by the suggestion[54] of a quadrupole polarization pattern in an anisotropic universe. Re-ionization can enhance this effect[55]. Polarization with a much smaller coherence angle will also arise due to inhomogeneity. The amplitude of the polarization pattern is of the order of the temperature anisotropy on a scale corresponding to the width of the last scattering shell[56]. If temperature anisotropy is ever detected, then polarization studies will provide an important test of the interpretation of such a result. It may be that temperature anisotropy searches will eventually be limited by confusion from faint sources. If this is the case, then polarimetry may provide the best way to view the primordial fluctuations.

Confidence intervals

The figures quoted above for upper limits on sky temperature fluctuations are a highly reduced form of the initial data, which typically consist of a set of N temperature difference estimates D_i together with estimates σ_i of the experimental uncertainty associated with each D_i. Before we consider the implications of the quoted upper limits it is important to understand the limitations of these statistics. To this end we will briefly discuss the method by which these confidence limits are arrived at and highlight some of the associated problems.

We will consider, for illustrative purposes, the simplified, though not entirely unrealistic example in which the D_i are measured in widely separated fields, and so can be considered to be statistically independent, and in which the instrumental variances are all equal: $\sigma_i^2 = \sigma_0^2$. In this case, the commonly adopted procedure[57] is to calculate the statistic $\chi^2 \equiv \Sigma D_i^2/\sigma_o^2$, and to ask, first, if there is firm evidence for fluctuations in excess of that produced by the receiver. Assuming that there is not, we then ask what is the value of σ^2_{ul} such that, if the true sky r.m.s. temperature difference is $\sigma_{\mathrm{sky}} = \sigma_{\mathrm{ul}}$, a value of χ^2 as low as that observed will occur only a fraction $1-c$ of the time. The number of σ^2_{ul} is our upper limit on the sky variance at confidence level c.

One problem with this method is that it requires a rather precise estimate of the instrumental variance σ_0 (ref. 58). For reasonably large N, the central limit theorem tells us that χ^2 will have a gaussian distribution with mean $N(\sigma_0^2+\sigma^2_{\mathrm{skv}})/\sigma_0^2$ and variance $2N(\sigma_0^2+\sigma^2_{\mathrm{sky}})^2/\sigma_0^2$. The upper limit is then $\sigma^2_{\mathrm{ul}} = \sigma_0^2\{(\chi_0^2/N)/(1-\alpha\sqrt{2}/N)-1\}$, where χ_0^2 is the value actually observed and α is the number of standard deviations associated with the confidence level c (for example, $\alpha = 1.65$ for $c = 95\%$). χ^2/N will be close to unity, $\chi^2/N = (1+\mathrm{O}(N^{-1/2}))$, and so $\sigma^2_{\mathrm{ul}} \approx \sigma_0^2 N^{-1/2}$. But the statistic χ^2 is inversely proportional to the estimate of σ_0^2, so σ^2_{ul} to be reliable, any systematic functional errors in σ_0^2 must be much smaller than $N^{-1/2}$. If we 'generously' overestimate σ_0^2, we will obtain much too low a value for σ^2_{ul} and we will be in danger of excluding the true hypothesis.

Another problem is that, by our definition of σ^2_{ul}, we will reject the true hypothesis a fraction $1-c$ of the time. Thus, if 20 upper limits at 95% confidence are found in the literature than the 'best' upper limit is likely to lie below the true value. The actual situation will be worse than this if there is any selection pressure to favour the appearance of 'good' results. This might occur if there is any tendency, either of a voluntary or involuntary nature, not to publish results which are deemed to be uninteresting compared to other published results. Such selection effects cause the number of independent samples of event space to exceed the number of published results. The rate of sampling of event space can also be boosted by combining sets of data from different runs, or if the decision to extend or curtail the experiment is made after any data have been collected.

If for these or other reasons, the value of σ^2_{ul} obtained has any influence on the publication of results, then the value of these constraints for rational decision making is debased. Ultimately, the only practical way to ensure that results are free of such bias is for the decision to publish or not to be made before the experiment is performed (with a blank space in the manuscript for σ^2_{ul} to be inserted, of course). It is after all only rational that the criterion for acceptance of such a paper should be the sensitivity of the apparatus, rather than the random number σ^2_{ul} which results at the end of the day.

Unreasonably restrictive upper limits may then occur by chance, or they may even be selected for. However, one should not simply accept this as a painful fact of life. Our procedure is guaranteed to reject the true hypothesis 5% of the time (for 95% confidence limits), but in any particular instance the data may give some indication as to whether this is such a case. For example, if $\sigma^2_{\mathrm{sky}} = 0$, then negative upper limits will occur. In such a case, we can be certain that we have been 'lucky'. There are other cases in which a positive limit results on the variance but in which one would be reluctant to reject the hypothesis that $\sigma^2_{\mathrm{sky}} \geqslant \sigma^2_{\mathrm{ul}}$. These are those cases in which, while our statistic indicates that the result is unlikely under the hypothesis $\sigma^2_{\mathrm{sky}} > \sigma^2_{\mathrm{ul}}$, the result is not significantly more likely under any alternative hypothesis. The problems associated with various statistical procedures for treating null results will be discussed elsewhere (A. N. Lazenby and N.K. in preparation), but for the present we note that the hallmark of such a spurious result is an unusually low value of χ^2. Visual inspection of the data presented in ref. 41 suggests this may have been the case in this instance. This is unfortunate because, taken at face value, this result rules out some interesting hypotheses.

Implications

Intermediate and large scales. The most immediate implication of these negative results is that we inhabit a universe that is currently uniform to a high degree of precision, at least on scales greater than the horizon size at last scattering. This inference is particularly certain because considerations of causality assure us that there is no mechanism which can render the Universe isotropic on scales exceeding the horizon scale at last scattering. Although we can observe only that region of space which is inside our present horizon, we can invoke the copernican principle to exclude such possibilities as that we inhabit an inhomogeneous universe which just happens to be isotropic about the Earth. We can also infer that the Universe is very close to homogeneous on scales much larger than the present horizon size[59]. If our universe is nearly homogeneous today, then it must have been prepared in a state with an extremely small amount of 'growing-mode' inhomogeneity present at early times; this preparation may have involved an inflationary phase[1]. The process which prepared the Universe may also have imprinted small departures from perfect homogeneity which could later grow to produce galaxies and other structure in the Universe. If this is the case, then it is most natural that the fluctuations should have curvature (specific binding energy) fluctuations which are independent of scale. Were this not the case, then it would be necessary to impose a special scale (that of curvature nonlinearity) in the initial conditions and this scale would not be many orders of magnitude different from the scale of galaxies. It is hard to see how the microphysics of the early Universe could produce such a scale.

On large scales the spectrum maintains its initial form. We can estimate the level of anisotropy which should be observed on intermediate and large angular scales, because the temperature fluctuations should be at least as large, to order of magnitude, as the specific binding energies of the largest structures we see today. Clusters have internal velocity dispersions of $v \approx 1000\ \mathrm{km\ s^{-1}}$ and so have specific binding energy $\phi \approx (v/c)^2 \approx 10^{-5}$. The binding energies of superclusters are thought to have similar magnitude and so we predict $\Delta T/T \approx 10^{-5}$. This is not much smaller than the current upper limits on these angular scales and so it seems that, with a modest improvement in the observational situation, we have the hope of either ruling out this picture or confirming the presence of these scale-free temperature fluctuations.

To obtain a quantitative constraint on $\Delta\rho/\rho$ on these scales, one can use the relation of Sachs and Wolfe[24] for the temperature fluctuation induced by a plane-wave density ripple. A convenient parameter to specify the amplitude of the perturbation is ϕ, the newtonian specific binding energy. The result of Sachs and Wolfe is simply that $\Delta T/T = \phi/3$. We have plotted the constraint on $(\Delta\rho/\rho)_\lambda$ using this relation, and the observed upper limits are shown in Figure 2. This constraint assumed that the initial fluctuation spectrum was initially isentropic.

Small scale. Observations on intermediate and large angular scales provide the best constraint on the amplitude of very-long-wavelength density fluctuations. On these scales, however, there is very little direct evidence for clustering of matter. On scales of $\leqslant 10h^{-1}$ Mpc we see strong clustering, at least of luminous material. The importance of small scale anisotropy studies is that they probe those length scales on which we actually see clustering, and so they provide a more direct test of theories of galaxy formation. Unfortunately, however, the interpretation of smallscale anisotropy is model-dependent: in order to apply this test one must first choose the nature of the initial fluctuations (for example, 'isothermal' or 'adiabatic'), assume some spectrum (such as a power law) and then propagate these fluctuations from some sufficiently early time until after the epoch at which the matter and radiation decouple. This linear evolution depends on the properties of the major constituents of the Universe at the time, so the abundances and masses of these particles must be specified. Up to this point the amplitude of the initial fluctuations is arbitrary. The next, and most uncertain, step is to normalize the calculation by requiring that the density fluctuations should have grown to the presently observed amplitude by today. One can then compare the predicted anisotropy with observations and thereby constrain the initial parameter space.

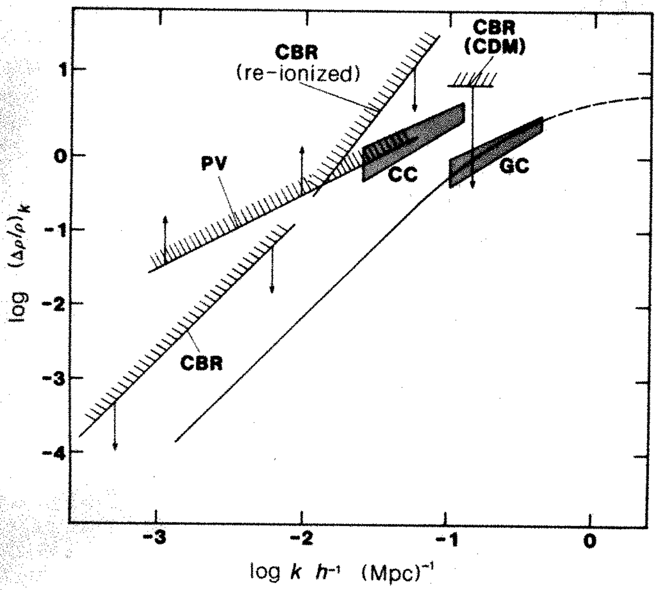


Fig. 2 Constraints on the power spectrum of density fluctuations in an Einstein–de Sitter ($\Omega = 1$) universe. $(\Delta\rho/\rho)_k \equiv k^{3/2}\delta_k/\sqrt{2}\pi$. The contribution to the r.m.s. density fluctuations per logarithmic interval of k, is plotted versus k, where $k(\text{Mpc}^{-1})$ is comoving wave number, $\delta_k = \int d^3k\, e^{ik.x}\,\Delta\rho(x)/\rho$, and x labels the spatial coordinates of an arbitrary point. Specifically, the mass fluctuation within a sphere of radius x is $\delta M/M = \int k^2\, dk|\delta_k|^2 W(kx)/\sqrt{2}\pi$, where the window function $W(kx) \approx 1$ for $kx \leqslant 1$ and $W(kx) \approx (kx)^{-4}$ for $kx \gg 1$. CBR, large- and intermediate-scale microwave anisotropy; we have taken $\Delta T/T \leqslant 3\times10^{-5}$ for $\theta \leqslant 10$ arc min and $\Delta T/T \leqslant 10^{-4}$ for $\theta \approx$ a few degrees. CBR(CDM), small-scale anisotropy limits in a universe dominated by cold dark matter. CBR(re-ionized), anisotropy limits if the intergalactic medium is re-ionized at large redshift and secondary fluctuations are generated as discussed in the text (see discussion of 'the re-ionized universe'). PV, constraint from peculiar velocity; power spectrum of galaxy clustering; CC, power spectrum of cluster clustering; CDM, theoretical preduction for density fluctuations in a cold dark matter model with $h = 0.5$. In a low-density universe ($\Omega = 0.2$), large- and intermediate-scale anisotropy limits are ~3 times lower; peculiar velocity limits are ~3 times higher; small-scale anisotropy limits are ~10 times lower.

In recent years, there have been interesting theoretical and observational developments concerning the small-scale anisotropy. On the theoretical side there has been a rise in popularity of models in which the bulk of the gravitating matter is in a form which is otherwise very weakly coupled to other forms of matter. These models generate a much smaller level of anisotropy than models in which the density is dominated by baryons. Meanwhile, the observational limits have improved, and now not only rule out baryon-dominated models (or at least those with 'adiabatic' fluctuations) but also place useful constraints on models involving weakly interacting 'dark' matter.

The normalization that has been used in these calculations is that the density contrast in spheres of radius $8h^{-1}$ Mpc should have unit variance today. This normalization is chosen because the fluctuations in galaxy counts in spheres of this size have unit variance and because it is thought that fluctuations on these scales should be described fairly accurately by linear theory, and so should be accessible to theoretical calculation. Ideally, it would be desirable to normalize on even larger scales where linear theory would be even more accurate, but unfortunately the estimates of fluctuations in galaxy counts on larger scales become very uncertain.

With this choice of normalization, the small-scale anisotropy limit allows one to exclude 'cold dark matter' models with $\Omega = 0.2$ (see the constraint labelled CBR (CDM) in Fig. 2). This conclusion is particularly interesting because this is the value for Ω implied if galaxies fairly trace the mass distribution. If the value of Ω is significantly greater than 0.2, then galaxies are, for some reason, more strongly clustered than the matter. If this is the case, then it is inconsistent to use the conventional normalization. Inclusion of this effect would lower the predicted $\Delta T/T$ still further in addition to the approximate Ω^{-1} scaling (by a factor of ~ 2)[60-62]; to detect such tiny fluctuations would require a great improvement over the current upper limits.

The possibility of ruling out a high-density Universe dominated by cold dark matter from small-angle anisotropy is therefore fairly remote. Large-angular-scale searches present a more hopeful prospect, although in this case the interpretation depends crucially on the assumed slope of the primordial spectrum of fluctuations. As we have mentioned, there are reasons to prefer the Zel'dovich, scale-free curvature spectrum. To rule out, or to confirm at the several-sigma level, the existence of the consequent scale-free temperature fluctuations would be a significant achievement.

Extrinsic anisotropy

The microwave sky has now been almost completely mapped at an angular resolution of a few degrees. The only convincing feature which has been observed to date is the 'dipole' anisotropy. Averaging together the two most precise of the recent measurements yields a dipole anisotropy with amplitude $\delta T/T = 1.2\times10^{-3}$. The usual interpretation of this anisotropy is that the Sun is moving at a velocity $v_{\text{solar}} \approx 380\ \text{km s}^{-1}$ towards an apex ~45° away from the Virgo cluster. The reason for this interpretation, rather than as a combination of extrinsic and intrinsic anisotropy, is that this dipole is at best an order of magnitude larger than the quadrupole or other low-order spherical harmonics on the microwave sky, and such an intrinsic pattern seems highly implausible. It may be possible to test this interpretation if the intrinsic quadrupole is really very small because, in addition to the dipole of amplitude v/c, our motion will also result[63] in an aligned quadrupole of amplitude v^2/c^2.

This measurement of the solar motion provides us with a constraint on large-scale density fluctuations because, if these induce an r.m.s. amplitude of $\gg 300\ \text{km s}^{-1}$, we must live in a very special position[64,65]. Unless there is some anthropic argument to favour the existence of observers in regions which seem to be special only in so far as they have an unusually low dipole anisotropy, we can exclude such hypotheses as unlikely. These hypothetical large-scale density fluctuations should be well described by linear theory, and we then have $v_{\text{rms}} \approx HL(\Delta\rho/\rho)_{\text{rms}}$, for fluctuations on length scale L. If we assume for the moment that the density fluctuations are a gaussian process, then the large-scale velocities will also be gaussian-distributed. We then have an upper limit at 95% confidence of $v_{\text{rms}} \leqslant 3v_{\text{solar}}$ as only ~5% of all observers have a velocity less than one third of the r.m.s. value. We have assumed a gaussian distribution, but this result is not very sensitive to the form of the tails of the actual distribution because we are calculating the probability of very small values of v/v_{rms}. The result would be misleading, however, if the probability $P(v) \neq v_{\text{rms}}^{-3}$ for $v \ll v_{\text{rms}}$, as would be the case if a small fraction of observers have $v \gg v_{\text{rms}}$ while most are nearly at rest. Note that we have ignored contributions to v_{solar} from small scales which are nonlinear ($\Delta\rho/\rho \gtrsim 1$). This is legitimate if we have no detailed knowledge of the form of these velocities, as the inclusion of small-scale $\Delta\rho/\rho$ will only broaden the distribution $P(v)$. But as we do have some knowledge of the Sun's motion in the Galaxy and the Galaxy's motion in the Local Group, we should incorporate this information. In doing so, we will be able to 'filter out' the nonlinear velocities and concentrate on those linear velocities of interest.

The first step in this process is to allow for the rotation of the Galaxy. This results in a larger velocity, $v_{\text{galaxy}} \approx 550\ \text{km s}^{-1}$, and therefore a weaker constraint than we would have if we had no knowledge of the galactic rotation. The reason for this is that the galactic rotation correction is almost parallel to the solar motion and is therefore unlikely (it means that we have one of the lowest dipole anisotropies of all observers in the Galaxy). However unlikely this state of affairs is *a priori*, it is hard to imagine that the galactic rotation estimate is wildly wrong, and so we must fully incorporate this information. The next step is to subtract the motion of the Galaxy with respect to other galaxies in the vicinity. Here, as we shall see, the situation is not so clear-cut. As before, the *a priori* expectation is that the velocity will decrease. However, as with the galactic rotation we should prepare ourselvles for possible surprises.

We determine the galactic motion with respect to a sample of galaxies around us by means of a 'Tully-Fisher' or 'Faber-Jackson' relation, or some analogous intrinsic correlation between a distance-dependent and a distance-independent quality. These 'L-σ' relations, or analogous relations, are established for a set of galaxies which we believe to be at the same distance from us (for example, in a cluster of galaxies). If we then assume that the correlation found is universal, this allows us to correct redshifts to true distances in a statistical manner by minimizing residuals. The validity of a velocity thereby obtained depends crucially on the assumed universality of the relation. If there are systematic differences in the zero-point of the relations from cluster to cluster, then these will be misinterpreted as peculiar velocities. An effect of this kind must be present at some level. In principle one can test for this by applying the method to clusters at a wide range of distances. Any spurious velocities will increase linearly with distance. These effects, if present at a significant level, do not however present any fundamental limit to the depth of sample for which one can determine the net velocity, as the number of clusters increases sufficiently quickly that the errors remain manageable.

The important question, for our purpose, is whether the source of our peculiar velocity has been localized within the sample. If so, we obviously get a stronger constraint on $\Delta\rho/\rho$ at large scales. It seemed, following the publication by Hart and Davies[66] that this was indeed the case. They studied a sample of galaxies with $v \leq 5{,}000\ \mathrm{km\ s^{-1}}$ using a radio variant of the Tully-Fisher (TF) technique, and found that the net motion of these galaxies with respect to the MBR was very small: $v_{\mathrm{HD}} = 130 \pm 70\ \mathrm{km\ s^{-1}}$. This result stands in contrast to the earlier result by Rubin *et al.*[67] for a sample of similar depth, although doubts about this result had already been expressed[68,69]. However, the Hart and Davies sample has now been reanalysed[70], and yields a much greater net motion of $\gtrsim 700\ \mathrm{km\ s^{-1}}$. An independent sample of similar depth was analysed by de Vaucouleurs and Peters[71], who found an intermediate value, $v_{\mathrm{dV}} \approx 300\ \mathrm{km\ s^{-1}}$. More recently, Collins *et al.*[72] and Burstein *et al.*[73,74] have found net streaming velocities of $\sim 700\ \mathrm{km\ s^{-1}}$ using slightly deeper samples, and are in approximate agreement with the original result of Rubin *et al.*[67]. However, these results had been preceded by those of Aaronson *et al.*[75], who used the infrared TF technique to analyse a sample of galaxies lying in the declination strip visible from Arecibo. They found a net motion of $\lesssim 200\ \mathrm{km\ s^{-1}}$, although the geometry of their sky coverage results in increased uncertainty for a velocity vector in the direction favoured by Burstein *et al.*[73,74] and by Collins *et al.*[72]. Another interesting result to emerge from the Aaronson *et al.*[75] sample is that the scatter of individual cluster velocities about the Hubble flow (or of any offsets in the TF relation interpreted as velocities) is very small. They obtained a limit of $\sim 200\ \mathrm{km\ s^{-1}}$ on the amplitude of such motions, roughly consistent with the individual cluster motions of $\sim 400\ \mathrm{km\ s^{-1}}$ (for a different sample) inferred by Burstein *et al.*[75], and suggesting that the infrared TF technique has very low intrinsic scatter.

The differences between the solutions found, both for net motion and for individual cluster motions, give some measure of the uncertainty associated with these methods. Certainly, the question of whether the source of the galactic motion has been localized remains in doubt. It is probably premature to try to use these data to constrain $\Delta\rho/\rho$ on very large scales. The most reliable datum we have is for the galactic motion of $\sim 550\ \mathrm{km\ s^{-1}}$ with respect to the microwave frame, and this results in the line labelled PV in Fig. 2. If a consistent picture for large streaming motions does emerge then this will imply, according to standard theory, sizeable density fluctuations ($\gtrsim 30\%$) on a scale of $\gtrsim 100h^{-1}$ Mpc. Such density fluctuations present a challenge to theorists: for instance, it appears that such fluctuations do not arise naturally in the cold dark matter model. A theory in which non-gaussian fluctuations emerge naturally, such as cosmic strings combined with hot dark matter, offers some prospect of simultaneously allowing large streaming motions and low $\delta T/T$. It should someday be possible to directly measure, or possibly exclude, such density fluctuations with large-scale redshift surveys.

Conclusions

The microwave background isotropy limits, as well as the measured dipole anisotropy and large-scale galaxy distribution, have already succeeded in severely constraining inflationary cosmological models. Indeed, the only survivors require an appeal to rather more contrived models than discussed here. One possibility requires a biasing scheme in which the observed galaxy distribution samples the gravitational effects of a universe with $\Omega = 0.1$, while the true Ω is unity[14,61]. Biasing schemes generally predict that $\sim 90\%$ of the Universe should contain 'failed galaxies', with the baryonic matter forming luminous galaxies only in the rare fluctuation peaks that sample $\sim 10\%$ of the Universe[76,77]. One alternative involves appeal to the vacuum energy density (or cosmological constant Λ): with $\Omega_{\mathrm{vac}} = 0.2$, where $\Lambda \equiv \Omega_{\mathrm{vac}}$, a combination of cosmological constant and cold dark matter can reduce both $\delta T/T$ and the large-scale peculiar velocity field to acceptable levels, while allowing light to be a good tracer of mass[78]. Another option is to let the cold dark matter (with $\Omega = 1$) decay into relativistic weakly interacting particles at a recent epoch (after galaxies have formed), thereby allowing astronomical measurements today to effectively sample a low-Ω universe[79,80]. This scheme apparently fails to produce flat galactic rotation curves[87] and generates an unacceptable feature in the galaxy correlation function[88], but it is consistent with $\delta T/T$ limits[81]. One could also revive hot dark matter in some non-baryonic guise such as the massive neutrino, with nonlinearity of the hot dark matter at a very recent ($z \approx 1$) epoch, as required by the various constraints such as $\xi(r)$ and the pairwise galaxy peculiar velocity distribution, but now seeding the Universe with large-amplitude small-scale fluctuations (cold dark matter, cosmic strings?) in order to allow some, if not all, galaxies to form at an earlier epoch[7,82,83]. Finally, the wakes produced by cosmic strings could themselves generate large sheet-like structures which might dominate the large-scale galaxy distribution[84–86]. String models naturally yield low (but ultimately observable) $\delta T/T$.

It is of course conceivable that we are completely on the wrong track. Perhaps the primordial fluctuation spectrum was not described by a scale-invariant power law, nor was it gaussian, nor even adiabatic. One might then resort to explosive amplification of primordial seeds to explain galaxy formation and leave little trace of the primordial initial conditions. This still poses the problem, however, of the origin of the large-scale clustering of the galaxies. We do observe large-scale correlations (superclusters and voids): surely these trace initial conditions. If our interpretation of the CBR in the standard Big Bang model is correct, and there is no real alternative, then we inevitably expect to see some residue of these initial conditions in the background radiation anisotropy. It would be too perverse of Nature to have thrown down an impenetrable screen which renders such fluctuations invisible. Eventual detection of $\delta T/T$ on some angular scale is inevitable, and it will surely elucidate one of the most challenging mysteries of the Big Bang theory, namely the origin and the formation of large-scale structure.

1. Guth, A. H. *Phil. Trans. R. Soc.* **A307**, 141-147 (1982).
2. Hawking, S. W. *Phys. Lett.* **115B**, 295-297 (1982).
3. Starobinskii, A. A. *Phys. Lett.* **117B**, 175-178 (1982).
4. Guth, A. & Pi, S.-Y. *Phys. Rev. Lett.* **49**, 1110-1113 (1982).
5. Bardeen, J. M., Steinhardt, P. J. & Turner, M. S. *Phys. Rev.* **D28**, 679-693 (1983).
6. Linde, A. *Phys. Lett.* **129B**, 177-181 (1983).
7. Silk, J. & Turner, M. *Phys. Rev.* **D.** (in the press).
8. Ostriker, J. P. & Cowie, L. *Astrophys. J.* **243**, L127-L132 (1981).
9. Ikeuchi, S. *Publs astr. Soc. Japan* **33**, 211-222 (1981).
10. Zel'dovich, Ya. B. *Mon. Not. R. astr. Soc.* **192**, 663-667 (1980).
11. Vilenkin, A. *Phys. Rep.* **121**, 263-315 (1985).
12. Wyse, R. G. & Jones, B. J. *Astr. Astrophys.* **149**, 144-150 (1985).
13. Blumenthal, G., Faber, S., Primack, J. & Rees, M. *Nature* **311**, 517-525 (1984).
14. Davis, M., Efstathiou, G., Frenck, C. S. & White, S. D. M. *Astrophys. J.* **292**, 371-394 (1985).
15. Frenk, C., White, S. D. M. & Davis, M. *Astrophys. J.* **271**, 417-430 (1983).
16. Efstathiou, G. & Silk, J. *Fundam. Cosmic Phys.* **9**, 1-138 (1983).
17. Bond, J. R. & Efstathiou, G. *Astrophys. J.* **285**, L45-L49 (1984).
18. Vittorio, N. & Silk, J. *Astrophys. J.* **285**, L39-L44 (1984).
19. Peebles, P. J. E. & Groth, E. *Astr. Astrophys.* **53**, 131-140 (1976).
20. Peebles, P. J. E. & Yu, J. T. *Astrophys. J.* **162**, 816-836 (1970).
21. Wilson, M. L. & Silk, J. *Astrophys. J.* **243**, 14-25 (1981).
22. Silk, J. *Nature* **215**, 1155-1156 (1967).
23. Silk, J. *Astrophys. J.* **151**, 459-471 (1968).
24. Sachs, R. K. & Wolfe, A. M. *Astrophys. J.* **147**, 73-90 (1967).
25. Peebles, P. J. E. *Astrophys. J.* L1-L5 (1982).
26. Bond, J. R. & Szalay, A. S. *Astrophys. J.* **274**, 443-468 (1983).
27. Zel'dovich, Ya. B. *Astr. Astrophys.* **5**, 84-89 (1970).
28. Efstathiou, G. & Bond, J. R. *Mon. Not. R. astr. Soc.* **218**, 103-121 (1986).
29. Suto, Y., Sato, K. & Kodama, H. *Astrophys. J.* **292**, L1-L4 (1985).
30. Fixen, D. J., Cheng, E. S. & Wilkinson, D. T. *Phys. Rev. Lett.* **50**, 620-622 (1983).
31. Kaiser, N. *Astrophys. J.* **282**, 374-381 (1984).
32. Hogan, C. *Mon. Not. R. astr. Soc.* **202**, 1101-1126 (1983).
33. Hogan, C. J. & Kaiser, N. *Astrophys. J.* **274**, 7-19 (1983).
34. Vilenkin, A. *Phys. Rep.* **131**, 263-315 (1985).
35. Kaiser, N. & Stebbins, A. *Nature* **310**, 391-393 (1984).
36. Gott, J. R. *Astrophys. J.* **288**, 422-427 (1985).
37. Brandenberger, R. H. & Turok, N. *Phys. Rev.* **D** (in the press).
38. de Lapparent, V., Geller, M. J. & Huchra, J. P. *Astrophys. J.* **302**, L1-L5 (1986).
39. Szalay, A., Bond, J. R. & Silk, J. in *Formation and Evolution of Galaxies and Large Structures in the Universe* (eds Audouze, J. & Tran Thanh Van, J.) 101-128 (Reidel, Dordrecht, 1984).
40. Hogan, C. J. *Astrophys. J.* **284**, L1-L4 (1984).
41. Uson, J. M. & Wilkinson, D. T. *Nature* **312**, 427-429 (1984).
42. Ostriker, J. & Vishniac, E. *Astrophys. J.* **306**, L51-L54 (1986).
43. Penzias, A. A. & Wilson, R. W. *Astrophys. J.* **142**, 419-421 (1965).
44. Melchiorri, F., Melchiorri, B., Ceccarelli, C. & Pietranera, L. *Astrophys. J.* **250**, L1-L4 (1981).
45. Fabbri, R., Guidi, I., Melchiorri, F. & Natale, V. *Phys. Rev. Lett.* **44**, 1563-1566 (1980).
46. Lubin, P., Villela, T., Epstein, G. & Smoot, G. *Astrophys. J.* **298**, L1-L5 (1985).
47. Smoot, G. F., Gorenstein, N. V. & Muller, R. A. *Phys. Rev. Lett.* **39**, 898-901 (1976).
48. Strukov, I. A. & Skulachev, D. P. *Soviet Astr. Lett.* **10**, 3-13 (1984).
49. Mandolesi, N. *et al. Nature* **319**, 751-753 (1986).
50. Fomalont, E. B., Kellermann, K. I. & Wall, J. V. *Astrophys. J.* **277**, L23-L26 (1984).
51. Kellerman, K. I., Fomalont, E. B., Weistrop, D. & Wall, J. *Highlights in Astronomy* (Reidel, Dordrecht, in the press).
52. Lubin, P. M. & Smoot, G. F. *Astrophys. J.* **245**, 1-17 (1981).
53. Lubin, P. M., Melese, P. & Smoot, G. F. *Astrophys. J.* **273**, L51-L54 (1983).
54. Rees, M. J. *Astrophys. J.* **153**, L1-L5 (1986).
55. Negroponte, J. & Silk, J. *Phys. Rev. Lett.* **44**, 1433-1437 (1980).
56. Hogan, C. J., Kaiser, N. & Rees, M. J. *Phil. Trans. R. Soc.* **A307**, 97-109 (1982).
57. Boynton, P. E. & Partridge, R. B. *Astrophys. J.* **181**, 243-253 (1973).
58. Lazenby, A. N. & Davies, R. D. *Mon. Not. R. astr. Soc.* **203**, 1137-1169 (1983).
59. Grishchuk, L. P. & Zel'dovich, Ya. B. *Sov. Astr. Lett.* **22**, 125-129 (1978).
60. Bardeen, J. M. in *Proc. Inner Space-Outer Space Symp.* (eds Kolb, E. W., Turner, M. S., Lindley, D., Olive, K. & Seckel, D.) 212-217 (University of Chicago Press, 1986).
61. Kaiser, N. in *Proc. Inner Space-Outer Space Symp.* (eds Kolb, E. W., Turner, M. S., Lindley, D., Olive, K. & Seckel, D.) 258-263 (University of Chicago Press, 1986).
62. Bardeen, J. M., Bond, J. R., Kaiser, N. & Szalay, A. *Astrophys. J.* **304**, 15-61 (1986).
63. Sunyaev, R. A. & Zel'dovich, Ya. B. *Mon. Not. R. astr. Soc.* **190**, 413-420 (1980).
64. Kaiser, N. *Astrophys. J.* **273**, L17-L20 (1983).
65. Vittorio, N. & Silk, J. *Astrophys. J.* **293**, L1-L4 (1985).
66. Hart, L. & Davies, R. D. *Nature* **297**, 191-196 (1982).
67. Rublin, V. C., Thonnard, N., Ford, W. K. & Robert, M. S. *Astr. J.* **81**, 719-737 (1976).
68. Fall, S. M. & Jones, B. J. T. *Nature* **262**, 547-460 (1976).
69. Schechter, P. L. *Astr. J.* **82**, 569-576 (1977).
70. Stavely-Smith, L. thesis, Univ. Manchester (1985).
71. de Vaucouleurs, G. & Peters, W. L. *Astrophys. J.* **287**, 1-16 (1984).
72. Collins, C. A., Joseph, R. D. & Robertson, N. A. *Nature* **320**, 506-508 (1986).
73. Burstein, D. *et al.* in *Galaxy Distances and Deviations From the Hubble Scale* (eds Madore, B. & Tully, R. B.) (Reidel, Dordrecht, in the press).
74. Lynden-Bell, D. *Q. Jl R. astr. Soc.* **27**, 319-327 (1986).
75. Aaronson, M. *et al. Astrophys. J.* **302**, 536-563 (1986).
76. Silk, J. *Astrophys. J.* **207**, 1-8 (1985).
77. Rees, M. J. *Mon. Not. R. astr. Soc.* **213**, 75P-81P (1985).
78. Peebles, P. J. E. *Astrophys. J.* **284**, 439-444 (1985).
79. Turner, M., Steigman, G. & Krauss, L. *Phys. Rev. Lett.* **52**, 2090-2093 (1984).
80. Gelmini, G., Schramm, D. & Valle, J. *Phys. Lett.* **146B**, 311-317 (1984).
81. Vittorio, N. & Silk, J. *Phys. Rev. Lett.* **54**, 2269-2272 (1985).
82. Achulli, S., Occhionero, F. & Scaramella, R. *Astrophys. J.* **299**, 577-582 (1985).
83. Umemusa, M. & Ikeuchi, S. *Astrophys. J.* **299**, 583-592 (1985).
84. Silk, J. & Vilenkin, A. *Phys. Rev. Lett.* **53**, 1700-1703 (1984).
85. Vachaspati, T. *Phys. Rev. Lett.* **57**, 1655-1657 (1986).
86. Brandenberger, R., Stebbins, A., Turok, N., Veeraraghavan, S. & Silk, J. *Astrophys. J.* (submitted).
87. Flores, R., Blumenthal, G. R., Dekel, A. & Primack, J. R., preprint SCIPP 86/64 (1986).
88. Veeraraghavan, S., Suto, Y. & Silk, J., *Astrophys. J.* (submitted).

ARTICLES

A molecular gradient in early *Drosophila* embryos and its role in specifying the body pattern

Paul M. Macdonald* & Gary Struhl*

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Ave, Cambridge, Massachusetts 02138, USA

After fertilization, the protein products of the Drosophila *homeobox gene* caudal (cad) *accumulate in a concentration gradient spanning the anteroposterior axis of the developing embryo. Mutations in the* cad *gene that reduce or eliminate the gradient cause abnormal zygotic expression of at least one segmentation gene* (fushi tarazu) *and alter the global body pattern.*

ONE recurring dilemma in developmental biology is the conflict between mosaic models of embryonic patterning which invoke rigid localizations of qualitatively distinct determinants, and regulative models, which postulate morphogen gradients or inductive cascades of cell-cell interactions. In the case of insect development, the classic experiments of Sander[1-4] and others[5-9] have ruled out mosaic models, except possibly for the segregation of the germ line[10], and argued strongly for anteroposterior gradients of morphogens that specify the body pattern. Further support for such a gradient model has come from the analysis of maternal effect mutations, such as *bicaudal*, *oskar* and *bicoid*, which alter the global segment pattern[11-18]. But concrete demonstrations of graded morphogens have been frustratingly absent.

During the past few years, molecular studies have led to the identification of a conserved structural domain, the homeobox, common to many genes which control segment development[19,20]. In general, homeobox containing genes have tightly restricted spatial roles[21-25]; hence, they are first activated in specific regions of the embryo around the time segments are established[19,26-33]. Recently, we (unpublished observations) and others[34,35] have identified a new homeobox gene, *caudal* (*cad*)[34], which has the unusual property of being expressed during oogenesis and early embryogenesis. Hence, it seemed possible that this atypical gene might have a more global role in organizing the segment pattern. Indeed, circumstantial evidence for such a role has been obtained by Mlodzik *et al.*[34] and Levine *et al.*[35] who observed that *cad* transcripts form an anteroposterior concentration gradient during the syncytial blastoderm stage when the segment pattern is being set up.

We have used antibodies directed against the *cad* protein, together with mutations in the *cad* gene, to examine its develop-

---

*Present address: Howard Hughes Medical Institute, Center for Neurobiology and Behaviour, Columbia University College of Physicians and Surgeons, 722 West 168 Street, New York, New York 10032.

mental role. We find that the *cad* protein is localized in nuclei and is distributed in an anteroposterior gradient throughout the embryo during most of the syncytial blastoderm stage. As observed for the *cad* messenger RNA[34,35], the *cad* protein gradient forms after fertilization, and disappears by the onset of gastrulation, when expression is restricted to a circumferential band of cells near the posterior pole. Further, we show that mutations which appear to inactivate the *cad* gene alter both segmentation and the expression of the pair-rule gene *fushi tarazu* (*ftz*)[36,26,27]. Thus we have directly visualized a graded gene product which acts in specifying the body pattern.

Establishment of a *cad* protein gradient

Rabbit polyclonal antibodies directed against the *cad* portion of a β-galactosidase-*cad* fusion protein (Fig. 1) were used to visualize the distribution of *cad* protein throughout embryogenesis as described (Figs 2 and 3).

During early *Drosophila* development, the single nucleus of the zygote undergoes a series of 13 largely synchronous divisions before cellularization (see refs 37, 38 for detailed descriptions; these divisions define fourteen stages numbered according to the nuclear division which completes each stage). Initially, the proliferating nuclei accumulate in the central, yolky portion of the egg. We detect little if any *cad* protein in the embryo during these early division cycles (stages 1-4); however, beginning in either stage 5 or 6, the cytoplasm and diffuse zones of 'structured cytoplasm'[38,39] around the nuclei stain faintly with the *cad* antibody (data not shown). During stage 7 most of the nuclei begin to migrate towards the periphery; nuclear migration continues during stage 8 and is complete by the end of stage 9 when peripheral nuclei at the posterior pole are incorporated into pole cells. In this period (stages 7-9), *cad* protein accumulates dramatically, especially in the posterior half of the body, forming a smooth concentration gradient along the anteroposterior axis (Fig. 2*a*; see also Fig. 2*b*). In addition, *cad* protein throughout the body becomes progressively more tightly localized to nuclei. The peripheral nuclei then undergo a further 4 rounds of syncytial divisions, during which time (stages 10-13), the concentration gradient persists and spans the length of the embryo (Fig. 2*b*, *c*, *d*). Note that graded nuclear expression of *cad* protein runs throughout the egg, including both peripheral and 'yolk' nuclei (Fig. 2*c*, *d*); posterior pole cell nuclei also stain intensely with the *cad* antibody.

Formation of the *cad* protein gradient does not depend either on nuclear division or zygotic gene expression. Unfertilized eggs collected within 90 minutes of laying show either no *cad* expression, or moderate, relatively uniform expression. In older unfertilized eggs, the levels of *cad* protein rise markedly, especially in the posterior portion of the egg, generating an anteroposterior gradient similar to that normally present in stage 9-13 embryos (data not shown).

Further polarization

Beginning in early stage 14, the pattern of *cad* protein distribution changes dramatically. Initially, antibody labelling subsides progressively from the anterior end of the embryo (Fig. 2*e*), while expression either persists or is enhanced in the posterior half. This results in a transient bipartite staining pattern in which the anterior 40% of the embryo appears devoid of *cad* protein, while relatively high levels of staining remain in the posterior 50%. The progressive loss of *cad* protein then continues into the posterior half of the embryo, extending more quickly along the ventral side (Fig. 2*f*, *g*). At the same time, *cad* protein levels decline in somatic nuclei at the posterior pole (between 0-13% egg length, EL, measured from the posterior pole) although not in the pole cells. These various changes in *cad* expression generate a diffuse posterior ring of *cad* protein which extends more anteriorly on the dorsal than the ventral side. Still later, as the blastoderm cellularizes and gastrulation ensues, this diffuse zone of expression sharpens into a discrete posterior

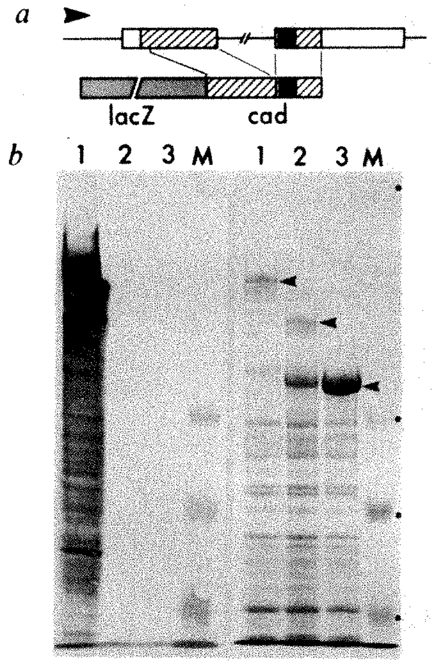


Fig. 1 Antibodies against the *cad* protein. *a*, Structure of the *cad* gene and derivation of a β-galactosidase-*cad* fusion protein. *b*, Specificity of the *cad* antibody. *a*, Zygotic and maternal transcripts of the *cad* gene are composed largely if not exclusively of two exons (wide bars) separated by an ~10 kilobase (kb) intervening sequence. Although the 5′ end of the first exon differs in the various maternal and zygotic transcripts, all have an identical long open reading frame. A 180-base-pair (bp) homeobox is positioned close to the splice junction in the second exon (ref. 34 and P.M.M., unpublished data; open boxes, untranslated 5′ and 3′ portions of the mature maternal transcript; hatched boxes, major open reading frame; black boxes, homeobox domain; arrowhead, direction of transcription). A *lacZ-cad* fusion gene (plc201) was constructed in the vector pUR290 (ref. 64) such that *cad* amino acids 29–425 (a total of 425 is predicted by the DNA sequence) were fused in frame to the carboxy terminus of β-galactosidase (the *cad* gene in p1c201 was derived almost entirely from genomic DNA except for a small segment covering the splice junction which was obtained from a cDNA clone; stippled box = the β-galactoasidase coding sequence). *b*, Total protein from IPTG-induced bacteria containing the *lacZ-cad* fusion gene plc201 (lane 1), a *lacZ-even-skipped* fusion gene including the *even-skipped* homeobox domain (lane 2) and the unfused *lacZ* gene from the fusion gene vector (pUR290) (lane 3) was separated by SDS-polyacrylamide gel electrophoresis, electroblotted onto nitrocellulose and probed with affinity-purified rabbit anti-*cad* antibodies prepared as described below (left panel). A duplicate gel stained with Coomassie Blue is shown in the right panel. No cross-reactivity to either β-galactosidase or homeodomain determinants is detectable. The heavy smear extending downward from the *cad* fusion protein bands in lane *C* is due to breakdown products of the unstable fusion protein. Size markers (highlighted by filled circles on the right) have relative molecular masses of 200,000 (M_r 200K) 97.4K, 68K and 43K.

Methods. The *cad* gene was cloned in a systematic screen[29] for new homeobox-containing genes, and the complete coding sequence of the gene determined directly (P.M.M., unpublished data). Positive confirmation that this gene is indeed the *cad* gene described by Mlodzik *et al.*[34] is as follows: (1) *in situ* hybridization experiments show that it is located in the 38E5,6 polytene band interval of salivary gland chromosomes, (2) it generates maternal and zygotic transcripts with sizes and temporal profiles similar to those described[34], (3) the DNA sequence is virtually identical to the partial *cad* cDNA sequence published previously[34]. A *lacZ-cad* fusion gene (plc201) was constructed as described in *a*. After IPTG induction of E. coli strain 71-18 (ref. 64) carrying plc201, the fusion protein was partially purified by preparative SDS-polyacrylamide gel electrophoresis, homogenized in Freunds complete adjuvant and injected subcutaneously at multiple sites in virgin female rabbits. Booster immunizations (in Freunds incomplete adjuvant) were given 2 months later, and serum collected after a further 10 days. Contaminating antibodies directed against β-galactosidase, other *E. coli* proteins, and generic homeodomain determinants were removed by successive passages over affinity columns prepared by coupling the desired ligand to Affigel 10 (Biorad). Ligands used were purified β-galactosidase (BRL), or a crude bacterial extract from cells containing a *lacZ-even-skipped* fusion gene (pleBB1, which includes the *even-skipped* homeobox). A similar column, prepared from the plc201 bacterial extract, was used for affinity purification of the anti-*cad* antibodies. Detailed descriptions of plasmid constructions and antibody purification are available upon request. The Western blot was prepared and stained by standard methods, using an NTB/BCIP alkaline phosphatase detection system (BRL).

ring 3-4 cells wide (Fig. 2*h*; a dorsal stripe of labelling sometimes persists transiently at around 50-60% EL). Note that throughout this process, the changes in *cad* distribution in the peripheral and yolk nuclei are tightly coordinated until the posterior ring is well defined, at which time the yolk nuclei stain poorly or not at all.

Expression of *cad* in subsequent development

At the onset of gastrulation, the presumptive mesoderm, including cells from the most ventral portion of the *cad* posterior ring, invaginate through the ventral furrow; soon thereafter, the pole cells, which continue to label with the *cad* antibody, begin migrating dorsally and anteriorly, surrounded by the remaining cells of the posterior ring. As the germ band extends, the pole cells and possibly some cells from the posterior ring appear to enter the posterior mid-gut invagination (Fig. 2*i*). Subsequently, cells in the developing posterior mid-gut and malpighian tubules (but not in the hindgut) express high levels of *cad* protein. The *cad* staining cells which remain on the outside of the embryo form a segment-length block of cells in the posterior germ band, which ultimately gives rise to parts of the larval anal pads (Fig. 2*j*). Based on double-labelling experiments with a monoclonal antibody directed against *Ultrabithorax* protein products[40], this block of cells appears to lie immediately posterior to parasegment 14 (see ref. 41) and may thus correspond to a cryptic fifteenth parasegment (data not shown).

Three additional features of *cad* expression are notable. First, in favourable preparations of embryos during germ band extension, we observe six narrow stripes of cells positioned at double segment length intervals (Fig. 2*i*). Second, a few bilaterally symmetric pairs of neurons are labelled in the neuromeres of parasegments 1-14 (ref. 42) after germ band shortening (later this labelling is retained in only two to four neuromeres in the thorax and anterior abdomen; Fig. 2*j*). In addition, a small cluster of neuroblasts perhaps belonging to a cryptic fifteenth parasegment labels at the posterior end of the nerve cord (data not shown). Third, persistent *cad* labelling is also observed during the remainder of larval development and in adults in the posterior mid-gut, malpighian tubules, and portions of the genital imaginal disc from which the adult analia may be derived (Fig. 2*j* and data not shown).

Mitotic cycle dependence

The large size and mitotic synchrony of the peripheral nuclei in the syncytial blastoderm stages allow us to monitor the presence and subnuclear distribution of the *cad* antigen throughout the mitotic cycle. As shown in Fig. 3*e*, the *cad* antibody does not label interphase nuclei uniformly, but rather reveals a punctate distribution of *cad* protein which is similar to, although not coincident with, that of AT rich satellite DNA (as revealed by the Hoechst stain which labels DNA in proportion to its AT-content[43]). In contrast, a monoclonal antibody, 8C5, which appears to recognize a ubiquitous chromosomal antigen[70] displays a more uniform and diffuse staining pattern (Fig. 3*f*).

At the onset of mitosis, nuclear *cad* labelling disappears almost completely (as previously noted for *engrailed* protein[44]), although both the Hoechst stain and the 8C5 antibody clearly label condensed chromosomes (Fig. 3*a-c*). Because the level of diffuse cytoplasmic staining seems to rise during this period, it is possible that the *cad* protein is actually displaced from condensed chromosomes rather than being rendered inaccessible to the antibody. By late telophase, nuclei begin to label again with the *cad* antibody, although punctuate staining is not evident until interphase.

These observations suggest that *cad* protein may be associated with discrete chromosomal sites during interphase (when transcription occurs) and that these sites may be closely linked or brought close together within the nucleus. Also, they suggest the possibility that *cad* protein may be displaced from the chromosomes during mitosis.

Mutations in *cad* alter segment specification

Four recessive mutations of the *cad* gene have been obtained using EMS mutagenesis ($cad^{1,2,3,4}$; see Fig. 4). Three of these four mutations ($cad^{2,3,4}$), as well as deletions[45] of the *cad* gene region, cause recessive lethality and eliminate detectable levels of *cad* protein in the diffuse posterior ring and later embryonic derivatives where *cad* is normally expressed (for example, Fig. 2*i, j*). They therefore behave as apparent null alleles and are referred to subsequently as cad^- mutations.

Homozygous cad^- embryos obtained from heterozygous mothers give rise to almost normal first instar larvae. However, these larvae lack cuticular structures of the posterior terminalia, most notably the anal tuft, parts of the anal pads, and the terminal sense organs[24] (sense organs 1 in ref. 46), all structures supposed to derive from a cryptic tenth abdominal segment[46] (or parasegment 15) (Fig. 4*d*). Thus embryos receiving maternally derived gene product, but lacking a functioning *cad* gene, form a normal segment pattern except for terminal structures derived from cells which display persistent zygotic *cad* expression.

When embryos derived from $cad^-/+$ parents are stained for expression of the *cad* gene product, all embryos before stage 14 show the characteristic patterns of *cad* protein, even though $\frac{1}{4}$ should be homozygous for the cad^- mutation. Thus, normal graded patterns of *cad* protein can be generated from maternally derived products of the *cad* gene, without any zygotic contribution. To examine the consequences of lacking *cad* gene function during this early period of embryogenesis, it is therefore neccessary to eliminate both the maternal and zygotic contributions. This has been done by generating clones of homozygous cad^- germ cells by gamma-ray induced mitotic recombination in $cad^-/+$ larvae (as in ref. 47), and then fertilizing eggs derived from such mutant germ line cells with cad^- sperm (see Fig. 5).

After fertilization of mutant cad^- eggs with sperm from $cad^-/+$ males, two classes of progeny were obtained in approximately equal numbers. The first class give rise to first instar larvae which usually hatched and often developed into normal fertile adults. However, many of these larvae were found to lack portions of the eighth abdominal segment (Fig. 4*c*); less frequently they also showed partial deletions of the fourth abdominal segment, and on occasion, of other, usually even-numbered, abdominal segments. These larvae almost certainly derived from mutant eggs fertilized by wild type sperm, because all of the progeny obtained from mutant eggs fertilized by wild type males developed in this fashion. Thus, even in the absence of maternally derived cad^+ function, zygotic expression from a single, paternally derived cad^+ gene is sufficient to allow normal, or almost normal, development.

Progeny of the second class show a dramatic alteration of the normal body pattern and presumably arise from fertilization of mutant eggs by mutant sperm (Fig. 4e-h). Although the segmentation phenotypes of these larvae are variable both from animal to animal, and between the left and right sides of single animals, certain common features emerge. First, the head and anterior thorax are usually normal, while the remainder of the body is severely shortened owing to variable deletions of many of the posterior segments. In general the second thoracic and odd-numbered abdominal segments are more resistant to deletion, so that most larvae show a phenotype resembling that of alternating segmental deletions caused by mutations in 'pair-rule' segmentation genes[48]. Second, most of the eighth abdominal segment and posterior terminalia (although curiously not some portions of the posterior spiracles) are absent and replaced by small plates of sclerotized cuticle resembling mouth-hooks. In addition small clusters of very fine ventral hairs (usually characteristic of the patch of ventral hairs in the middle of the prothorax) lie between these plates and the preceding

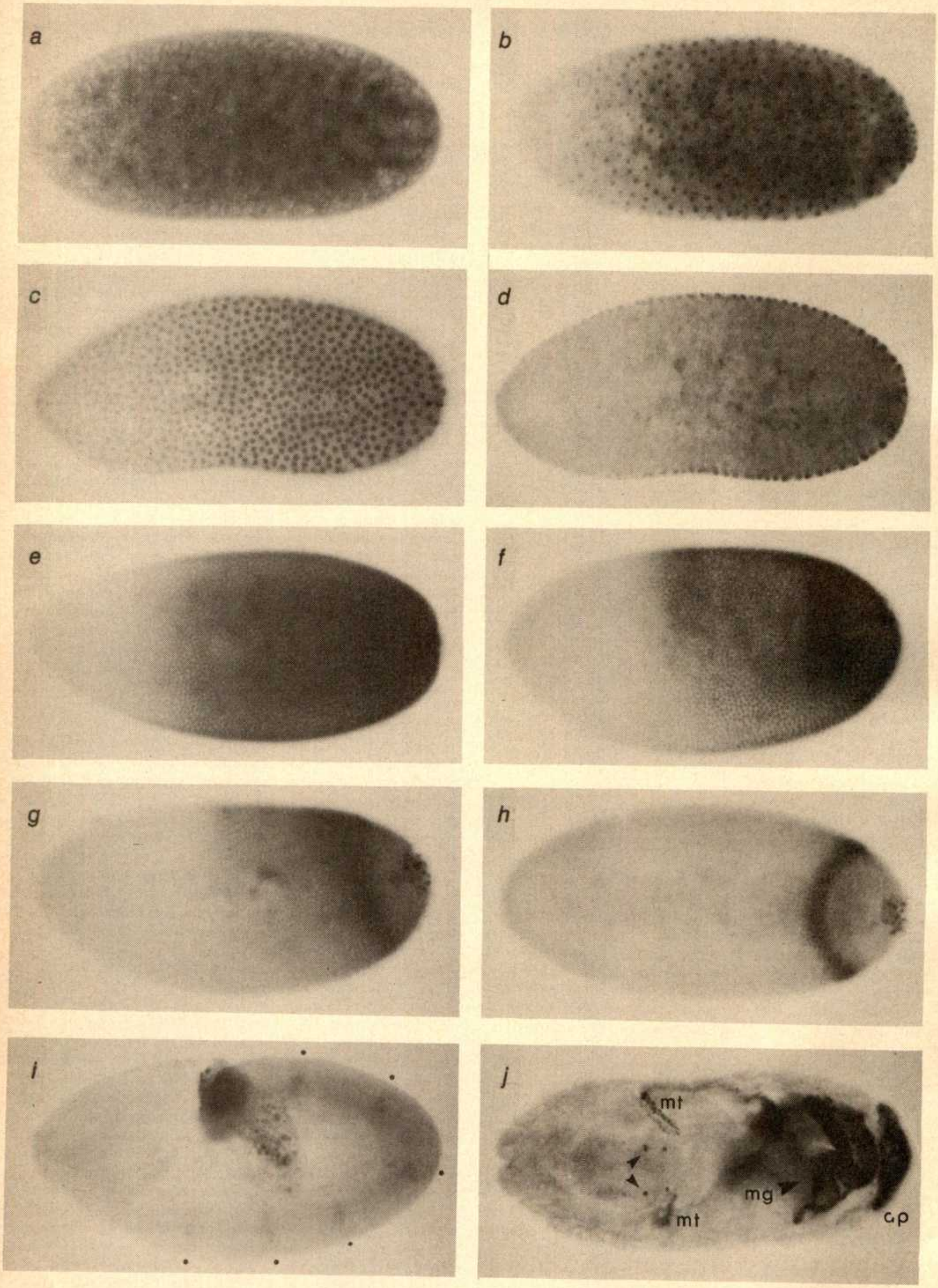


Fig. 2 (legend opposite).

Fig. 2 Immunohistochemical localization of *cad* protein during embryogenesis. All embryos are oriented with anterior to the left and the plane of focus through the middle of the embryo (*a, d, i, j*) or at the periphery (*b, c, e, f, g, h*); embryos before gastrulation are staged according to the nuclear division cycle (for instance, stage 9 falls between the 8th and 9th nuclear divisions). *a*, Stage 8 embryo. All of the nuclei are still located within the central yolky region of the egg. Staining with the *cad* antibody is most intense in the nuclei, but note that islands of cytoplasmic staining surround each nucleus. Graded distribution of the *cad* protein is already apparent, as displayed by the gradual reduction of staining towards the anterior pole. *b*, Stage 10; Nuclear migration to the periphery is complete and the pole cells have budded off. The *cad* protein gradient now extends along the entire body axis, peaking at the posterior pole. This pole to pole concentration gradient is maintained until the beginning of stage 14. Note that it is unlikely to be due to either a fixation artifact or unequal accessibility of the antibody, as both the DNA dye Hoechst 33258 as well as a monoclonal antibody directed against a chromosomal antigen reveal uniform staining throughout this period in triple labelling immunofluorescent experiments (see Fig. 3); the close to uniform distribution of *cad* protein observed in *BicD* embryos (Fig. 5a) also argues against such possibilities. *c, d*, Stage 12 (periphery and inside respectively). The *cad* gradient is confined predominantly to nuclei and is apparent both in peripheral and yolk nuclei. Pole cell nuclei, just visible in *d*, stain as intensely as the neighboring somatic nuclei. *e*, Early stage 14. During stage 14 the distribution of *cad* protein undergoes a dramatic change, becoming progressively more polarized until it is restricted to a single posterior ring 3–4 cells wide at the onset of gastrulation. This embryo shows the early phase of the redistribution in which *cad* protein is removed progressively from the anterior pole. *f*, Early-mid stage 14 (dorsal side at top). Progressive loss of *cad* staining continues extending more quickly along the ventral surface; staining also begins to decline in somatic nuclei between 0–13% egg length (EL; from the posterior pole). Thus, *cad* staining now appears predominantly in a diffuse posterior ring which extends dorsally to about 55% EL. *g*, Mid-stage 14 (dorsal side at top). *cad* staining continues to decline from the anterior and dorsal directions, although intense staining persists in the posterior ring (now somewhat narrower) and in the pole cells. Note that the anterior dorsal boundary of *cad* labelling remains quite distinct, despite the falling level of expression. Note also that here, as in *f*, *cad* expression in the yolk nuclei (out of focus, but visible as dark spots in the center of the embryo) parallels that of the peripheral nuclei, declining progressively in the anterior and ventral portions of the egg. *h*, Late stage 14 (dorsal surface). By the onset of gastrulation, *cad* protein is found in a tightly restricted posterior ring (which is ~3–4 cells wide around the entire embryo) and in the pole cell nuclei. In some preparations, a relatively faint dorsal stripe 1–2 cells wide can be seen anterior to the posterior ring. *i*, Germ band extension (lateral aspect). As the germ band extends dorsally and anteriorly, most *cad* expressing cells from the posterior ring remain at the terminus of the germ band, though some may enter the posterior midgut invagination together with the labelled pole cells. Transient *cad* expression is detected in rows of cells (●) which transect the elongating germ band, spaced at approximately double segment intervals from the block of *cad* staining cells at the end of the germ band. *j*, 15 h embryo (ventral aspect). The *cad* staining can be seen in cells giving rise to portions of the anal pads (*ap*), the posterior midgut (*mg*), and the malpighian tubules (*mt*) which wander in and out of the plane of focus. Staining is also apparent in one or more pairs of bilaterally symmetric nuclei in several neuromeres (arrow heads). Similar neuroblast staining is seen in most if not all of the neuromeres at an earlier stage (after germ band shortening), but is restricted to only a few thoracic or anterior abdominal neuromeres as embryogenesis continues.
Methods. Embryos aged for 0–15 h after fertilization were dechorionated, fixed, devitellinezed, and incubated with the primary antibody as described previously[65,27,44]. The embryos were washed under standard conditions, and incubated with a goat anti-rabbit IgG secondary antibody coupled to biotin (Vector Labs) which had been preadsorbed with 0–12 h embryos. The embryos were then washed again and incubated with a streptavidin-horseradish peroxidase conjugate (BRL). After further washing, the histochemical stain was developed with hydrogen peroxide and diamino-benzidine to yield a permanent signal. Stained embryos were dehydrated in 100% ethanol and mounted directly in a mixture of ~1.5 g ml^{-1} dried canada balsam in methyl salicylate[66]. Preadsorption of the *cad* antibodies with the *cad* affinity resin (see Fig. 1 legend) eliminated the staining patterns shown here.

abdominal segment (usually A7). Note that, despite the severity of the segmental deletions, the overall polarity of cuticular derivatives such as ventral and dorsal hairs is usually unaffected.

Thus, complete absence of the *cad* gene function dramatically alters the normal segment pattern. In parallel with the polarized distributions of *cad* gene product during the syncytial blastoderm stage, anterior head and thoracic segments are usually unaffected, whereas more posterior portions of the body show progressively more severe defects.

Maternal versus zygotic contributions

The extreme segmental defects displayed by *cad*$^-$ mutant embryos can be rescued either by maternal or zygotic activity alone. Consequently, common features of maternal and zygotic

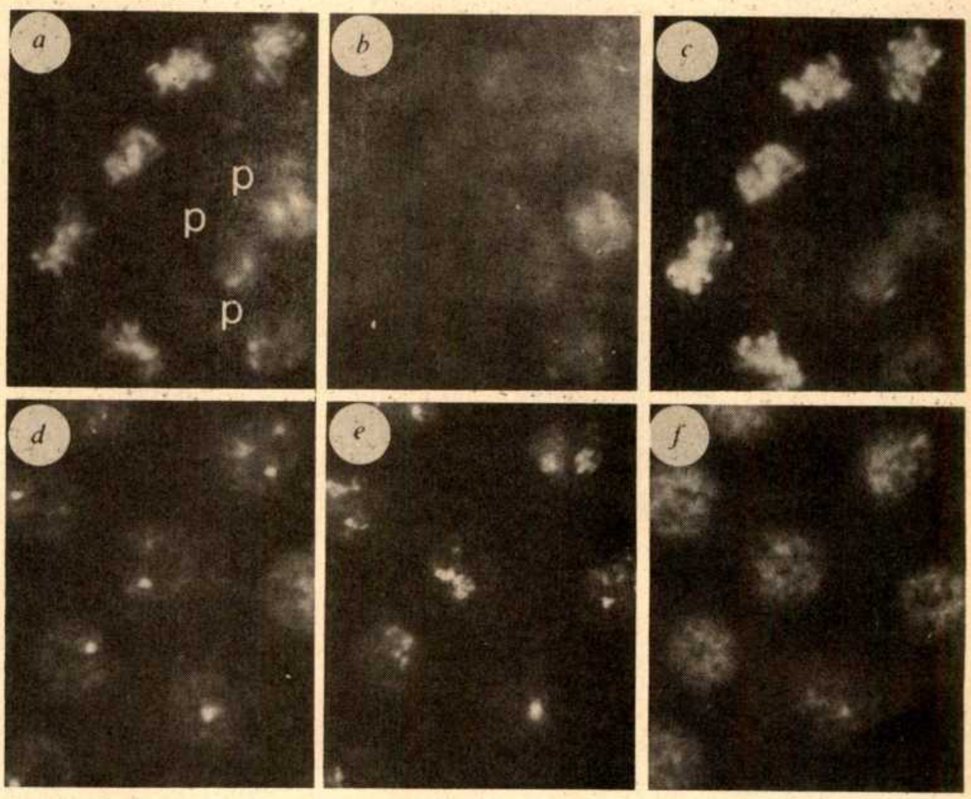


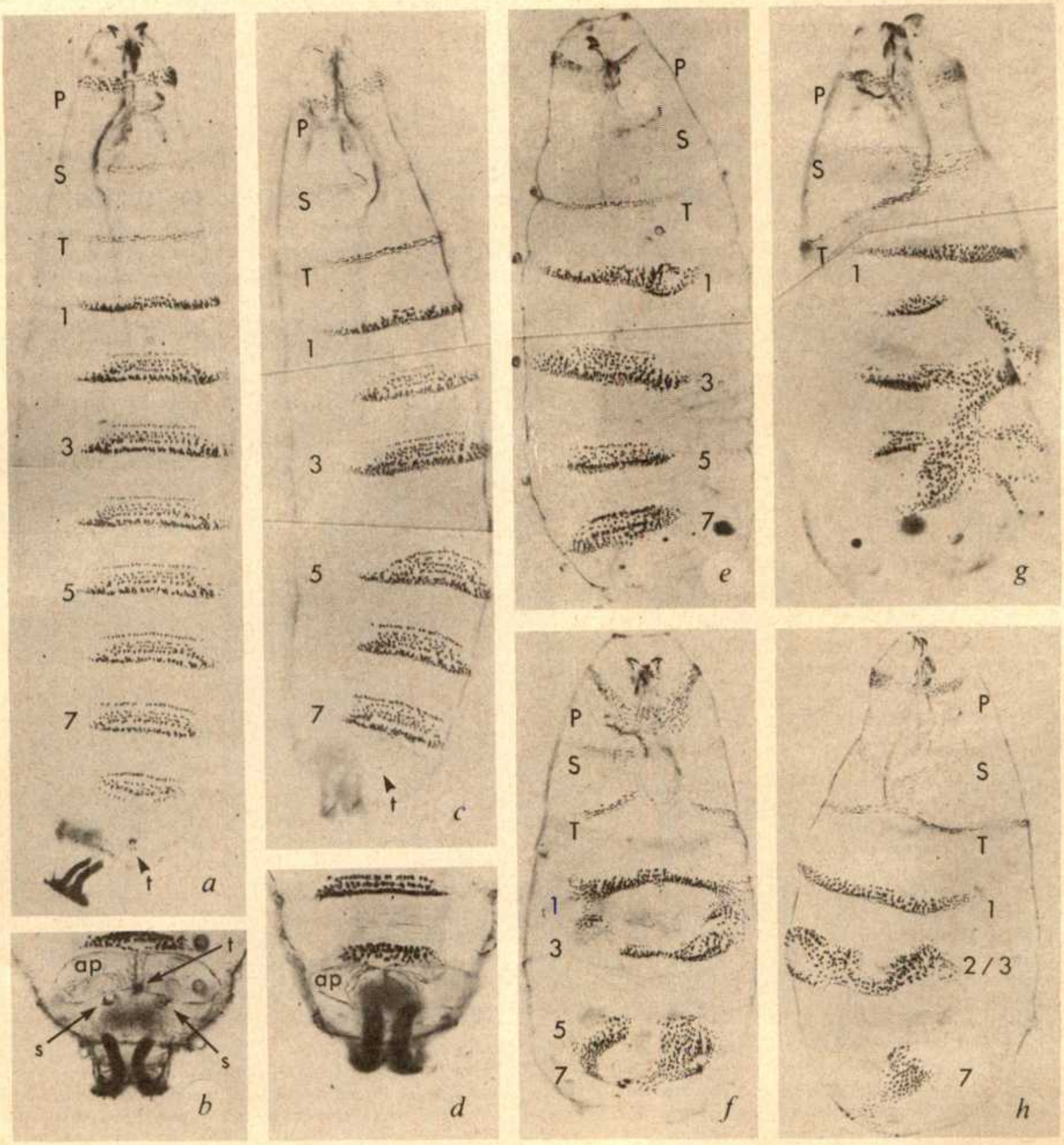
Fig. 3 Mitotic cycle dependence and subnuclear distribution of *cad* staining. The relative distributions of DNA, *cad* protein, and a chromosomal antigen recognized by the monoclonal antibody 8C5 (ref. 70) are shown during the 12th mitotic division (*a*, *b* and *c*, respectively) and during the interphase preceding this division (*d*, *e* and *f*). In both bases, the nuclei were located close to the posterior pole (indeed, in *a*, *b* and *c* the interphase nuclei of three pole cells (*p*) are adjacent to the mitosing syncytial nuclei). Note that the Hoechst stain (*a*) and 8C5 antibody (*c*) clearly show condensed chromosomes in mitosing peripheral nuclei, whereas only faint labelling is detected with the *cad* antibody (*b*). In contrast, interphase peripheral nuclei (*d*, *e*, *f*) as well as interphase pole cell nuclei (*a*, *b*, *c*) show punctate staining with both the Hoechst stain (*a*, *d*) and *cad* antibody (*b*, *e*) and a more diffuse pattern of staining with the 8C5 antibody (*c*, *f*). Note that the punctate pattern of the Hoechst stain (presumably labelling AT-rich satellite sequences[43]) does not coincide with that revealed by the *cad* antibody.
Methods. Embryos were prepared as in Fig. 2 and probed with the rabbit anti-*cad* antibody as well as the mouse monoclonal antibody 8C5. After washing, the embryos were incubated with two secondary antibodies, a preadsorbed goat anti-rabbit immunoglobulin G (IgG) coupled to biotin (Vector) and a goat anti-mouse IgG coupled to rhodamine (Tago Labs). The embryos were then washed again and incubated with a fluorescein-avidin conjugate (Vector Labs), stained with the Hoechst dye, washed, mounted in 50% glycerol with 2.5 mg ml^{-1} p-phenylenediamine added as an anti-bleaching agent, and examined with epi-fluorescence illumination.

protein expression may define critical aspects of *cad* function which are responsible for ensuring normal segmentation. As described above, all embryos derived from *cad*$^-$/+ embryos show normal graded patterns of *cad* protein until the beginning of stage 14. But after this time ~$\frac{1}{4}$ fail to show the normal bipartite pattern of expression or the subsequent polarization which gives rise to the posterior ring. Instead, embryos in this novel class display a gradually fading gradient of *cad* protein, although pole cells continue to stain intensely well into gastrulation (Fig. 5*d*; compare with Fig. 2*g*). Paternally rescued embryos (*cad*$^-$/+ embryos derived by fertilizing *cad*$^-$ eggs with wild type sperm) resemble maternally rescued *cad*$^-$ embryos in forming an anteroposterior gradient of *cad* protein, although the levels of *cad* expression are atypically low, close to the limit of detection (data not shown). In stage 14, the level of *cad* antigen rises significantly in posterior portions of paternally rescued embryos, forming the normal bipartite pattern (Fig. 5*e*) and then becoming progressively restricted to the posterior ring (as in Fig. 2*e*–*h*; note that no *cad* protein is expressed in pole cells).

Thus, the common link between the maternal and zygotic contributions to *cad* protein expression appears to be an anteroposterior differential in protein concentration which arises during the syncytial blastoderm stage. This can be generated solely by gradients of maternally derived gene product, or by graded and then more polarized patterns of zygotic expression.

Symmetrical *cad* expression in embryos

If the gradient of *cad* protein plays an instructive role in specifying the body pattern, mutations which cause global changes in the body plan should alter segmentation and the distribution of *cad* protein in a tightly coordinated fashion. Dominant mutations of the *BicD* locus act during oogenesis to alter the axial polarity of the early embryo[15]. Specifically, embryos derived from mutant mothers display a 'double abdomen' phenotype in which the body is composed of two sets of posterior abdominal segments and terminalia arranged in mirror symmetry[15]. When embryos from mothers transheterozygous for two dominant *BicD* mutations (see Fig. 5, legend) are stained with the *cad* antibody, they show corresponding alterations in the normal patterns of *cad* expression. First, in most embryos, *cad* antigen is distributed almost uniformly in nuclei throughout the body until early stage 14 (Fig. 5*a*). The intensity of labelling then



declines in the central portion of the embryo as well as in narrow zones at both ends (Fig. 5*b*). Finally, these embryos display a stable pattern of mirror symmetric posterior rings (Fig. 5*c*). These results show that *BicD* mutations cause a reorganization of the *cad* protein distribution which parallels that of the body pattern.

Expression of *fushi tarazu* in *cad*⁻ embryos

The tendency of *cad*⁻ mutant embryos to show pair-rule segmentation defects in the posterior half of the body suggests that one role of early *cad* protein activity may be to coordinate periodic expression of pair-rule genes[48] along the anteroposterior body axis. We therefore examined the expression pattern of the pair-rule gene *fushi tarazu* (*ftz*)[36,26,27] in cad⁻ embryos using a polyclonal antibody directed against the *ftz* protein (ref. 27). In wild type embryos, the *ftz* protein is expressed in 7 stripes of blastoderm nuclei shortly before the onset of gastrulation: the first six stripes are of equal width (3-4 nuclei) whereas the posterior seventh stripe is wider (5-7 nuclei)[27]. When *cad*⁻ eggs were fertilized by sperm from *cad*⁻/+ males, and the resulting embryos stained for *ftz* protein expression at the onset of gastrulation, two classes of progeny were observed in similar numbers. Embryos of the first class almost certainly correspond to paternally rescued *cad*⁻/+ embryos and approximate wild type *ftz* expression, although stripes 2 and 4 are slightly wider, and stripes 3, 5 and 6 slightly narrower than normal (Fig. 6*a*). In contrast, embryos of the second class almost certainly correspond to homozygous *cad*⁻ embryos and show highly variable and abnormal patterns of *ftz* expression, stripes 2 and especially 4 being much wider than in wild type embryos, and stripes 3, 5, 6 and 7 being much narrower or absent (Fig. 6*b*). Thus, absence of *cad* protein activity early in embryogenesis causes abnormal periodic expression of the *ftz* protein.

Discussion

Both embryological[1-9,49-52] and genetic[11-18] experiments provide compelling evidence that the segment pattern of the insect body is specified by a dynamic system of spatial cues generated early in development. Moreover, they have led to explicit proposals that this system consists of one or more morphogen gradients organized by the spreading influences of determinants at the anterior and posterior poles[2,17,18,53]. Here, we visualize a protein gradient which arises during the period when the body plan is established. Moreover, we show that the formation of this gradient depends on maternal factors such as the *BicD* gene[15] which influence the localization of anterior and posterior polar determinants[17,18]. Finally, we demonstrate that activity of the gradient protein is critical for setting up the normal segment pattern. Thus, we believe we have identified a graded morphogen which responds to early spatial cues and participates in organizing the body plan.

Role of the *cad* protein gradient

The developing pattern of *cad* protein expression is complex, involving first the accumulation of *cad* protein in the form of a continuous anteroposterior gradient, followed by progressive loss of the protein from the anterior pole until it is restricted to a ring of cells close to the posterior pole. Hence, the morphogenic role of the *cad* gene product could depend on the concentration gradient *per se*, or more simply, on the consequent generation of an anteroposterior differential of *cad* protein activity. Two lines of evidence support the view that graded *cad* expression *per se* is required for normal segmentation. First, embryos lacking the *cad* function show a spectrum of graded segmental defects which appear to correlate with the graded expression of *cad* protein during the syncytial blastoderm stage. Hence, the requirement for *cad* function appears proportional to the relative levels of protein along the body. Second, early graded expression of maternally derived gene product is sufficient to rescue the development of *cad*⁻ zygotes (except for derivatives of the posterior terminalia which require persistent zygotic expression). Neither line of evidence is compelling, however. Indeed, we also observe that zygotic expression, which generates first an abnormally shallow gradient, and then a more normal bipartite distribution can largely rescue the *cad*⁻ mutant phenotype. This last result argues against attaching a critical significance to the absolute concentration of *cad* protein along

Fig. 4 Segmental phenotypes resulting from loss of *cad* gene function. Four recessive mutations of the *cad* gene (*cad*^{1,2,3,4}) were obtained after EMS mutagenesis and the phenotypes of embryos lacking either maternal gene function (*c*), zygotic gene function (*d*), or both (*e–h*) have been determined as described below. Micrographs show the ventral cuticle patterns of first instar larvae (dark field illumination); the ventral dentical bands associated with the pro- (*P*), meso- (*S*) and metathorax (*T*) and the odd numbered abdominal segments (1, 3, 5, 7) are labelled; t, anal tuft (a patch of ventral denticals positioned just posterior to the anal pads); *a*, wild type; *b*, wild type posterior terminalia. Note the bilaterally symmetric anal pads (ap), the anal tuft (t) and the terminal sense organs (s) which lie just posterior and to either side of the tuft. *c*, paternally rescued *cad*^{2}/+ larva derived from a homozygous *cad*^{2} oocyte. Note the absence of the dentical bands associated with abdominal segments 4 and 8; neither segment is entirely deleted, as dorsal structures from both segments (not visible here) remain. *d*, Posterior portion of a homozygous *cad*^{3} larva derived from *cad*^{3}/+ parents. Larvae of this genotype are generally normal except for the absence of the anal tuft and portions of the anal pads. The terminal sense organs also appear to be absent or malformed. *e–h*, *cad*^{2}/*cad*^{3} larvae derived from homozygous *cad*^{2} oocytes. In general, embryos of this genotype usually form normal heads, part or all of the prothorax, and then a spectrum of progressively more extreme segmental deletions along the anteroposterior axis. As described in the text, an 'even-skipped' pair-rule modulation is often apparent (as in *e* and *f*), though more extreme deletions (*h*) or left-right differences in segmental deletions (*g*) can obscure this aspect of the phenotype. The polarity of cuticular structures within each segment is usually normal, though local perturbations are sometimes present in embryos such as *g*.

Methods. The *cad* gene maps to the polytene band interval 38E5,6 as determined by *in situ* hybridization experiments (ref. 34 and our unpublished data), a region which is deleted by a number of well characterized chromosomal deficiencies[45,67]. Two of these deficiencies, *Df(2L)TW1* and *Df(2L)TW2*, overlap, deleting a common region containing polytene bands 38A7-B1-38E6-E9 (ref. 45); a third deficiency, *Df(2L)Sd*^{R68} (37A/B-38E2-E5; ref. 67), deletes much of this common region, but not the 38E5,6 interval (we have confirmed these chromosomal assignments by determining that the *Df(2L)Sd*^{R68} chromosome, but not the *Df(2L)TW1* or *Df(2L)TW2* chromosomes can generate *cad* gene transcripts (P.M.M., unpublished data). We therefore performed a standard F2 screen of ~4,000 chromosomes for EMS-induced mutations which cause partial or complete lethality when *in trans* to *Df(2L)TW2*. All mutations were then retested against the remaining two deletions to eliminate most of the non-*cad* mutations. Of the remaining 30 mutations, 4 fall into a single complementation group which appears to define the *cad* gene by the following criteria: (1) all four mutations specifically cause loss of larval or adult structures which derive from embryonic or larval cells showing persistent zygotic expression of the *cad* protein; (2) approximately $\frac{1}{4}$ of the embryos generated by flies heterozygous for the *cad*^{2}, *cad*^{3} or *cad*^{4} mutations show no *cad* protein attributable to zygotic expression of the *cad* gene (as is observed for progeny from *Df(2L)TW1*/+ and *Df(2L)TW2*/+ parents), and (3) *cad*^{2}/+ embryos derived from *cad*^{2}/*cad*^{2} oocytes express *cad* protein during the syncytial blastoderm stage and in all subsequent stages, whereas *cad*^{2}/*cad*^{3} embryos derived from *cad*^{2}/*cad*^{2} oocytes do not express *cad* antigen at any time (*cad*^{1} homo- or hemizygotes express *cad* antigen, but are also able to survive to adulthood, generating male and female flies lacking anal plates; consequently, this mutation is likely to reduce rather than eliminate *cad* gene function). Finally note that although the three apparent null mutations of *cad* eliminate detectable *cad* protein expression, they may nevertheless retain low levels of *cad* protein activity and hence not be true amorphic mutations. Germ line clones of homozygous *cad*^{2} and *cad*^{3} cells were generated by standard means (see, for example ref. 47) using a dominant female sterile mutation, *Fs(2)1* (courtesy of Janos Szabad and Trudi Schupbach). Briefly, *b pr cad*^{2}/*Fs(2)1* (or *b pr cad*^{3}/*Fs(2)1*) larvae were γ-irradiated during the first instar and the resulting adult females crossed to *b pr cad*^{3}/*CyO, pr* (or *b pr cad*^{2}/*CyO, pr*) males. *Fs(2)1*/+ female germ cells either give rise to woefully abnormal eggs or no eggs at all; however, homozygous *Fs(2)1*$^+$ germ cells resulting from mitotic recombination proximal to the *Fs(2)1* mutation form normal eggs. Individual females laying normal eggs were selected and their progeny examined. Most (~90%) of these females generated the two classes of progeny described in the text, indicating that most recombination events that occur proximal to the *Fs(2)1* gene also occur proximal to the *cad* gene. Moreover, the paternally rescued (*b pr cad*⁻/*CyO, pr*) progeny from these females often survived to adulthood and were invariably purple in phenotype, confirming that they resulted from proximal mitotic recombination events (the remaining few females laying normal eggs generally gave rise to curly-winged progeny which had phentoypically wild type or purple eyes, indicating that mitotic recombination had occurred proximal to the *Fs(2)1* gene, but distal to the tightly linked *cad* and *pr* genes). Mutant larvae were dechorionated and dissected, when necessary, and mounted directly in a 1:1 mixture of lactic acid and Hoyer's mountant[68,69].

the body, and suggests that a coarse anteroposterior differential in *cad* function is required for normal segmentation.

Irrespective of which component of early *cad* expression is responsible for its morphogenic function, it is plausible that *cad* serves as a molecular link between more primary spatial determinants and the localized expression of zygotic control functions. For example, graded or differential expression of *cad* protein might be responsible for regulating the spatial expression of particular gap genes such as *hunchback*, *Kruppel* and *knirps*[46] which would then establish the segment pattern via the correct activation of pair-rule and homeotic genes[54-56]. Alternatively, graded *cad* expression might directly influence pair-rule and homeotic gene functions, some of which normally display anteroposterior gradients in their spatial expression[28,29,57,58]. Our demonstration that *cad*$^{-}$ embryos alter both segmentation and the periodic expression of the pair-rule gene *ftz* over a large region of the body provide direct support for the view that *cad* protein regulates the expression of pair-rule and homeotic genes either directly or via its effects on particular gap genes. The presence of a homeobox domain within the *cad* protein and our finding that *cad* protein is preferentially localized in transcriptionally active interphase nuclei both support this interpretation.

Generation of the *cad* protein gradient

If, as we propose, the dynamic pattern of *cad* expression is a response to the influence of more primary polar determinants, the mechanism by which the *cad* protein gradient forms is likely to be of considerable interest. One striking feature of our results is that *cad* protein expression appears to be graded both in peripheral and yolk nuclei and shifts posteriorly in these two populations of nuclei in a tightly coordinated fashion. Thus, as suggested previously by the initial pattern of activation of the segmentation gene *Kruppel*[59], the generation and interpretation of graded spatial information in insect embryos need not be a special attribute of the peripheral egg cortex, but may occur uniformly throughout the egg cell.

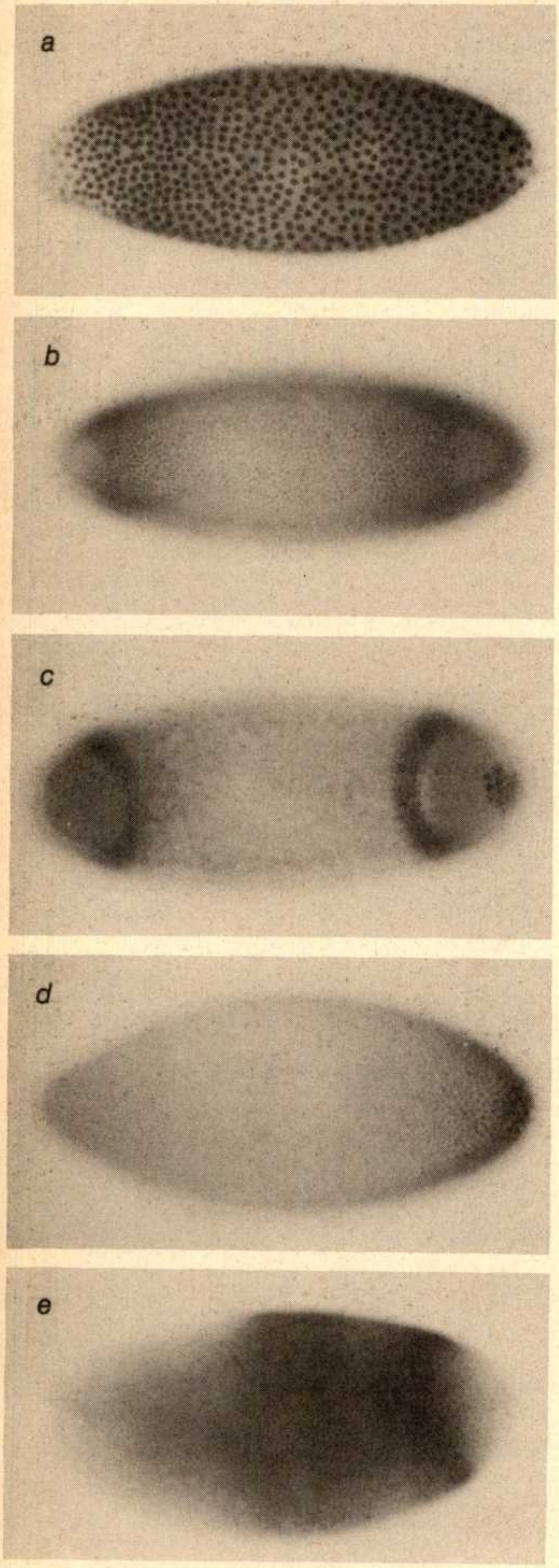


Fig. 5 Distribution of *cad* protein in *BicD* and *cad*$^{-}$ mutants. All embryos are oriented with anterior to the left and the plane of focus at the periphery of the egg. *a*, Stage 12 *BicD* embryo; *cad* protein is most abundant near the posterior pole, declines slightly towards the middle of the embryo and remains at a constant high level into the anterior region. The light staining at the anterior tip is commonly observed. Some variation in the overall pattern is seen in a population of embryos, as expected from the varying segmental phenotypes seen in populations of mutant embryos. *b*, Mid-stage 14 *BicD* embryo. Expression in the posterior ring, including the refinement of the initially broad region of staining, occurs roughly as normal, except that the ring forms at both ends. The pole cell nuclei (which are not in this plane of focus, but see *c*) stain intensely, as normal; there are no pole cells (or equivalent staining) at the anterior pole. *c*, Late stage 14 *BicD* embryo (dorsal aspect). Posterior and anterior rings are fully formed, both 3-4 nuclei wide. The width of the anterior ring can vary considerably, sometimes forming a stripe 10 or more cells wide. *d*, Mid-stage 14 *cad*$^{-}$ embryo derived from a *cad*$^{-}$/+ oocyte. The last traces of the gradient of *cad* protein can be seen near the labelled pole cell nuclei (see Fig. 2*g* for a wild type embryo at a comparable stage). *e*, Early-mid stage 14 embryo derived from a *cad*^{2} oocyte fertilized by wild type sperm (dorsal aspect). *cad* protein expression is similar to a comparably staged wild type embryo (Fig. 2*f*) except that no staining is observed in the pole cells. **Methods.** Antibody staining was as described (Fig. 2, legend), except for *e*; in this case, a strepavidin-alkaline phosphatase conjugate was used in place of the strepavidin-horseradish peroxidase conjugate, and the histochemical stain generated with the *black* kit (Vector Labs). *BicD* embryos were obtained from mothers transheterozygous for two different *BicD* alleles, *BicD*$^{71.34}$ and *BicD*$^{\mathrm{IIIF.48}}$ (ref. 15). Embryos that were *cad*$^{-}$ were obtained as $\frac{1}{4}$ of the progeny of a cross between mothers heterozygous for *Df(2L)TW2* and fathers heterozygous for *Df(2L)TW1*; both deficiencies delete the *cad* gene[45]. Similar results were obtained using embryos from selfcrosses of balanced *cad*$^{2-4}$ mutants. The *cad*^{2} eggs fertilized by wild type sperm were obtained by fertilizing females carrying germ line clones of the *cad*^{2} mutation with *cad*$^{+}$ males (see Fig. 4).

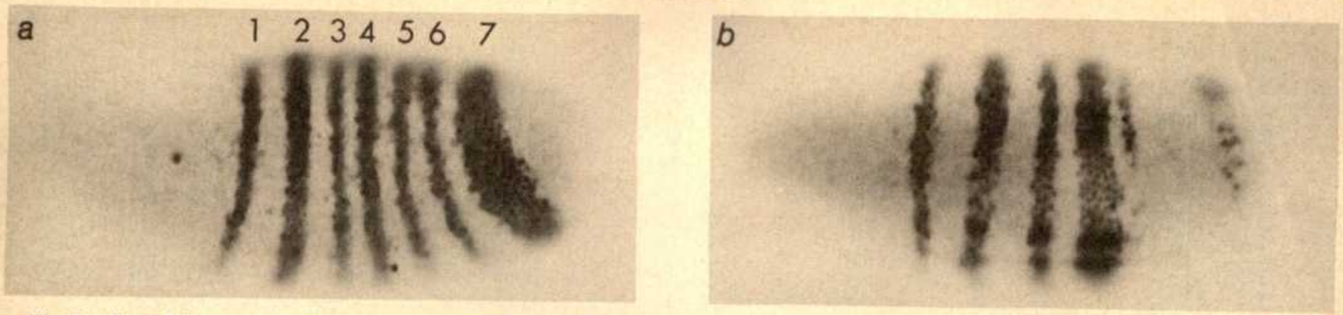


Fig. 6 Altered *ftz* protein expression in paternally rescued and cad^- mutant embryos; $cad^2/+$ (*a*) and cad^2/cad^3 (*b*) embryos derived from homozygous cad^2 oocytes, stained for *ftz* protein expression at the onset of gastrulation (lateral aspect, anterior to the left, dorsal at the top). Wild type embryos at this stage normally express six stripes of *ftz* protein of equal width followed by a broader seventh stripe. In paternally rescued $cad^-/+$ embryos (*a*), stripes 2 and 4 are slightly wider, and stripes 3, 5 and 6 are slightly narrower than in wild type. This tendancy is dramatically accentuated in cad^- embryos (*b*), leading to the loss of stripe 6 and severe reduction of stripes 5 and 7. Immunohistochemical staining as in Fig. 4*e*, except that the NBT/BCIP detection system (BRL) was used instead of the Vector Labs *black* kit.

Although uniformly high levels of *cad* RNA are found throughout the early embryo[34,35], we detect *cad* protein only after the fourth or fifth nuclear division. Soon thereafter, as the *cad* protein gradient forms, the distribution of *cad* RNA also becomes graded. This correlation suggests that a single process is responsible for localizing both the mRNA and protein. Because graded *cad* protein expression can be generated in unfertilized eggs as well as in cad^- embryos derived from heterozygous mothers, this process must depend on post-ranscriptional events, rather than on graded zygotic transcription.

In principle, formation of the *cad* protein gradient might reflect a passive response to changes in the distribution of the RNA (for example, mobilization of the transcripts, or local degradation by a specific, spatially restricted nuclease). Alternatively, both the RNA and protein gradients might be controlled at the translational level. One example of how this could occur is through differential initiation of translation along the anteroposterior body axis. Assuming that mRNA associated with ribosomes is shielded from nonspecific nucleases[60,61], efficiently initiated translation near the posterior pole would produce a high level of *cad* protein and at the same time stabilize the RNA, while infrequently translated *cad* RNA at the anterior pole would be prone to degradation. Analysis of the *cad* gene (P.M.M., unpublished data) provides suggestive evidence that supports a translational control model, as the various maternal and zygotic transcripts include sets of in-frame start and stop codons upstream from the probable initiator AUG codon; similar elements clearly act in translational control of the yeast GCN4 mRNA[62,63]. Polar determinants, such as those posited by Sander[2] and others[9,17,18], might regulate translation of *cad* mRNA either directly, or indirectly through their effects on other control molecules.

We thank Barry Ganetzky, Trudi Schupbach, Janos Szabad, Eric Wieschaus, and Ted Wright for fly stocks, Seymour Benzer for the monoclonal antibody 8C5, Sean Carroll and Matthew Scott for the *ftz* antibody, Sean Carroll, George Deegan, Tad Goto, and Ariel Ruiz i Altaba for advice or assistance in preparation of antibodies, Anne Ephrussi, Janice Fischer, Peter Lawrence, Christianne Nüsslein-Volhard and Robin Wharton for critical comments on the manuscript, and Tom Maniatis for advice and support. This work was supported by the U.S. National Institute of Health (Program Project Grant 1PO1GM29301 to Tom Maniatis), the Howard Hughes Medical Institute, and the Damon Runyon-Walter Winchell Cancer Fund (Postdoctoral fellowship DRG-768 to P.M.M.).

Received 22 September; accepted 10 November 1986.

1. Sander, K. *Wilhelm Roux Arch. dev. Biol.* **151,** 430–497 (1959).
2. Sander, K. *Wilhelm Roux Arch. dev. Biol.* **151,** 660–707 (1960).
3. Sander, K. in *Symposium on the Germ Cells and Earliest Stages of Development* (ed. Pallanza, M.) 338–353 (Institut International d'Embryololgie, 1961).
4. Herth, W. & Sander, K. *Wilhelm Roux Arch. dev. Biol.* **172,** 1–27 (1973).
5. Sander, K. *Adv. Insect Physiol.* **12,** 125–238 (1976).
6. Schubiger, G. & Wood, W. J. *Am. Zool.* **17,** 565–576 (1977).
7. Schubiger, G., Moseley, R. C. & Wood, W. J. *Proc. natn. Acad. Sci. U.S.A.* **74,** 2050–2053 (1977).
8. Vogel, O. *Wilhelm Roux Arch. dev. Biol.* **182,** 9–32 (1977).
9. Frohnhöfer, H. G., Lehmann, R. & Nusslein-Volhard, C. *J. Embryol. exp. Morphol.* (in the press).
10. Illmensee, K. & Mahowald, A. P. *Proc. natn. Acad. Sci. U.S.A.* **71,** 1016–1020 (1974).
11. Nüsslein-Volhard, C. *Wilhelm Roux Arch. dev. Biol.* **183,** 249–268 (1977).
12. Nüsslein-Volhard, C. in *Determinants of Spatial Organisation* (eds Subtelney, S. & Konigsberg, I. R.) 185–211 (Academic, New York, 1979).
13. Bull, A. *J. exp. Zool.* **161,** 221–242 (1966).
14. Lohs-Schardin, M. *Wilhelm Roux Arch. dev. Biol.* **191,** 28–36 (1982).
15. Mohler, J. & Wieschaus, E. F. *Genetics* **112,** 803–822 (1986).
16. Schüpbach, T. & Wieschaus, E. *Wilhelm Roux Arch. dev. Biol.* **195,** 302–317 (1986).
17. Lehman, R. & Nüsslein-Volhard, C. *Cell* **47,** 141–152 (1986).
18. Frohnhöfer, H. G. & Nüsslein-Volhard, C. *Nature* **324,** 120–126 (1986).
19. McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. & Gehring, W. J. *Nature* **308,** 428–433 (1984).
20. Scott, M. P. & Weiner, A. J. *Proc. natn. Acad. Sci. U.S.A.* **81,** 4115–4119 (1984).
21. Lewis, E. B. *Nature* **276,** 565–570 (1978).
22. Duncan, I. & Lewis, E. B. in *Developmental Order: Its Origin and Regulation, 40th Symposium of the Society for Developmental Biology* (ed. Subtelny, S.) 533–554 (Liss, New York, 1982).
23. Struhl, G. *Nature* **293,** 36–41 (1981).
24. Struhl, G. *J. Embryol. exp. Morphol.* **76,** 297–331 (1983).
25. Struhl, G. *Nature* **318,** 677–680 (1985).
26. Hafen, E., Kuroiwa, A. & Gehring, W. J. *Cell* **37,** 833–841 (1984).
27. Carroll, S. B. & Scott, M. P. *Cell* **43,** 47–57 (1985).
28. Harding, K., Rushlow, C., Doyle, H. J., Hoey, T. & Levine, M. *Science* **233,** 953–959 (1986).
29. Macdonald, P. M., Ingham, P. & Struhl, G. *Cell* (in the press).
30. Akam, M. *EMBO J.* **2,** 2075–2084 (1983).
31. Akam, M. E. & Martinez-Arias, A. *EMBO J.* **4,** 1689–1700 (1985).
32. Levine, M., Hafen, E., Garber, R. L. & Gehring, W. J. *EMBO J.* **2,** 2037–2046 (1983).
33. Harding, K., Wedeen, C., McGinnis, W. & Levine, M. *Science* **229,** 1236–1242 (1985).
34. Mlodzik, M., Fjose, A. & Gehring, W. J. *EMBO J.* **4,** 2961–2969 (1985).
35. Levine, M. *et al. Cold Spring Harb. Symp. Quant. Biol.* **50,** 209–222 (1985).
36. Wakimoto, B. T. & Kaufman, T. C. *Devl Biol.* **81,** 51–64 (1981).
37. Zalokar, M. & Erk, I. *J. Microsc. Biol. Cell.* **25,** 97–106 (1976).
38. Foe, V. E. & Alberts, B. M. *J. Cell Sci.* **61,** 31–70 (1983).
39. Karr, T. L. & Alberts, B. M. *J. Cell Biol.* **102,** 1494–1509 (1986).
40. White, F. A. H. & Wilcox, M. *Cell* **39,** 163–171 (1984).
41. Martinez-Arias, A. & Lawrence, P. A. *Nature* **313,** 639–642 (1985).
42. Hartenstein, V. & Campos-Ortega, J. A. *Wilhelm Roux's Arch. dev. Biol.* **193,** 308–325 (1984).
43. Weisblum, B. & Haenssler, E. *Chromosoma* **46,** 255–260 (1974).
44. DiNardo, S., Kuner, J. M., Theis, J. & O'Farrell, P. H. *Cell* **43,** 59–69 (1985).
45. Wright, T. R. F., Hodgetts, R. B. & Sherald, A. F. *Genetics* **84,** 267–285 (1976).
46. Sato, T. & Denell, R. E. *Devl Biol.* **116,** 78–91 (1986).
47. Wieschaus, E. & Szabad, J. *Devl Biol.* **68,** 29–46 (1979).
48. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287,** 795–801 (1980).
49. Yajima, H. *J. Embryol. exp. Morphol.* **8,** 198–215 (1960).
50. Yajima, H. *J. Embryol. exp. Morphol.* **12,** 89–100 (1964).
51. Kalthoff, K. *Wilhelm Roux Arch. dev. Biol.* **168,** 63–84 (1971).
52. Kalthoff, K. *Photochem. Photobiol.* **18,** 355–364 (1973).
53. Meinhardt, H. *Models for Biological Pattern Formation* (Academic, London, 1982).
54. Ingham, P., Ish-Horowicz, D. & Howard, K. *EMBO J.* **5,** 1659–1665 (1986).
55. Carroll, S. B. & Scott, M. P. *Cell* **45,** 113–126 (1986).
56. Meinhardt, H. *J. Cell Sci.* (1986).
57. Kilchherr, F., Baumgartner, S., Bopp, D., Frei, E. & Noll, M. *Nature* **321,** 493–499 (1986).
58. Weir, M. P. & Kornberg, T. *Nature* **318,** 433–439 (1985).
59. Kniple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A. & Jackle, H. *Nature* **317,** 40–44 (1985).
60. Zitomer, R. S., Montgomery, D. L., Nichols, D. L. & Hall, B. D. *Proc. natn. Acad. Sci. U.S.A.* **76,** 3627–3631.
61. Losson, R. & Lacroute, F. *Proc. natn. Acad. Sci. U.S.A.* **76,** 5134–5137 (1979).
62. Mueller, P. P. & Hinnebusch, A. G. *Cell* **45,** 201–207 (1986).
63. Tzamarias, D., Alexandraki, D. and Thireos, G. *Proc. natn. Acad. Sci. U.S.A.* **83,** 4849–4853 (1986).
64. Ruther, U. & Muller-Hill, B. *EMBO J.* **2,** 1791–1794 (1983).
65. Mitchison, T. J. & Sedat, J. *Devl Biol.* **99,** 261–264 (1983).
66. Struhl, G. *Devl Biol.* **84,** 372–385 (1981).
67. Brittnacher, J. G. & Ganetzky, B. *Genetics* **103,** 659–673 (1983).
68. Van der Meer, J. M. *Drosophila Inf. Serv.* **52,** 160.
69. Struhl, G. *Nature* **308,** 454–457 (1984).
70. Fujita, S. C., Zipursky, S. L., Benzer, S., Ferrus, A. & Shotwell, S. L. *Proc. natn Acad. Sci. U.S.A.* **79,** 7929–7933 (1982).

Infrared variability of the BL lacertae object OJ287 since its outburst in 1983

W. K. Gear, E. I. Robson & L. M. J. Brown

School of Physics and Astronomy, Lancashire Polytechnic,
Preston PR1 2TQ, UK

In early 1983, OJ287 was seen to undergo an outburst in its optical and infrared emission[1]. Since then optical monitoring studies[2–4] have shown a general decline in the source brightness, with considerable fluctuations. We have monitored the near-infrared emission since the outburst. The fluxes have fluctuated considerably, with the lowest recordings being an order of magnitude less than those measured during the outburst. We have found an excellent correlation between infrared flux and spectral index; as the source gets fainter the spectrum gets steeper, and vice versa. We interpret this in terms of outbursts being due to injection or reacceleration of electrons with a constant energy index which subsequently steepens as the electrons are affected by radiation losses.

We began observing OJ287 in February 1983 as part of a larger programme studying the infrared to millimetre variability properties of a sample of Blazars[5,6]. Observations have been made at J (1.25 μm), H (1.65 μm), K (2.20 μm) and L (3.45 μm) on 16 different nights over the period February 1983 to April 1986. All data were obtained from the United Kingdom Infrared Telescope (UKIRT), using common-user InSb photometers. Calibration was performed against stars from the observatory standards list.

The results of the observations and calibration procedure are shown in Table 1. In Table 2, we also list the infrared spectral indices, obtained by a least-squares fit method, for the 14 of these occasions on which we had complete JHKL coverage. We also checked each spectrum for departures from a power law. Given the rather small wavelength range of our data, and taking into account the uncertainties, we found no significant departure from a single power law at any epoch. Figure 1 plots the variability of the infrared emission over the period 1983-86, where to avoid overcrowding, we have averaged together spectra taken over a period of <1 week. Note that none of the spectra in Fig. 1 has been displaced; the different fluxes represent genuine variability. The extent of variability shown in Fig. 1 is such that to be sure of defining the overall continuum of OJ287, one would need to obtain the whole spectrum over no more than a single night, and preferably an even shorter time than this.

Table 1 Infrared observations of OJ287

	J	H	K	L
3 February 1983	31.0±0.5	42.2±0.6	59.4±1.2	101.6±3.6
5 February 1983	34.2±0.4	47.6±0.3	64.2±0.3	101.6±3.3
6 March 1983	38.5±1.3	51.4±3.4	69.0±3.9	119.0±10.0
29 January 1984	14.7±0.2	22.0±0.2	29.6±0.2	53.2±1.0
7 February 1984	11.8±0.7	16.5±1.1	22.5±1.7	41.7±3.8
8 February 1984	12.4±0.6	17.4±0.9	25.0±1.3	42.8±2.0
8 March 1984	26.7±1.3	35.9±1.8	47.5±2.4	82.3±4.1
18 April 1984	6.2±0.3	8.8±0.4	12.5±0.7	22.4±1.1
19 April 1984	5.9±0.1	8.5±0.1	11.8±0.1	—
21 April 1984	7.0±0.1	10.4±0.3	13.8±0.2	27.2±1.8
28 November 1984	3.0±0.1	4.6±0.2	7.1±0.2	14.6±1.3
22 December 1984	4.0±0.2	5.9±0.3	9.6±0.5	17.2±0.9
22 February 1985	10.6±0.5	13.9±0.7	20.0±1.0	32.4±1.6
24 April 1985	16.5±0.2	22.4±1.1	29.8±1.5	—
19 February 1986	5.5±0.2	7.8±0.3	11.9±0.4	25.3±2.5
24 February 1986	4.7±0.2	7.8±0.3	11.6±0 4	23.5±0.8

Table 2 Near-infrared spectral indices

Date	
3 February 1983	−1.16±0.03
5 February 1983	−1.08±0.02
6 March 1983	−1.09±0.08
29 January 1984	−1.23±0.02
7 February 1984	−1.23±0.10
8 February 1984	−1.22±0.06
8 March 1984	−1.11±0.07
18 April 1984	−1.27±0.06
21 April 1984	−1.23±0.03
28 November 1984	−1.54±0.07
22 December 1984	−1.48±0.07
22 February 1985	−1.12±0.06
19 February 1986	−1.43±0.07
25 February 1986	−1.56±0.05

Because of the extreme variability of OJ287 we cannot claim to have fully sampled its temporal behaviour; however we can search for overall patterns of behaviour, which are evidence for a single origin for the emission over this wavelength range, and can lead to insights into the physical mechanism producing the variability. Even from Fig. 1 it can be seen that there is a correlation between flux and spectral index. To clarify this, Fig. 2 shows the JHKL spectral indices plotted against the log of the J-band flux. For these 14 data, performing a linear regression taking into account the uncertainty in both the flux and spectral index, we find a correlation coefficient of 0.84, representing >99% certainty of correlation. The correlation is in the sense that as the flux decreases the spectrum steepens. Conversely, as the flux increases, the spectrum flattens.

Also plotted in Fig. 2, are data from ref. 1, taken at the peak of the 1983 outburst, and refs 7, 8, taken in January to April 1983. Combining these with the present data, we find an even better correlation coefficient of 0.91. This well-defined correlation confirms the trend found for a sample of Blazars including OJ287, but with less data[6].

The power-law infrared continuum, extreme variability and high polarization[1] observed in OJ287 indicate a synchrotron origin of the infrared emission. We can explain the behaviour observed in terms of a recurrent injection or reacceleration

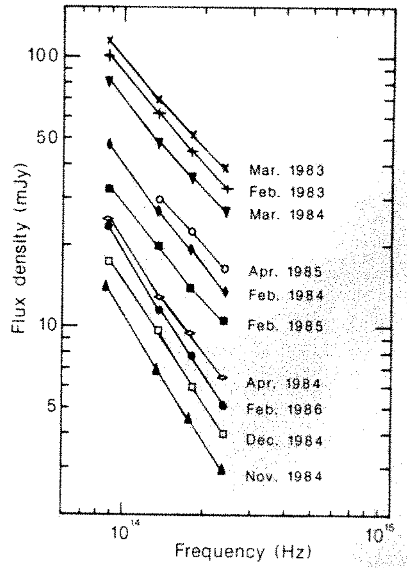


Fig. 1 The infrared spectrum of OJ287 at nine epochs over the period 1983-86. Note that none of the spectra has been displaced.

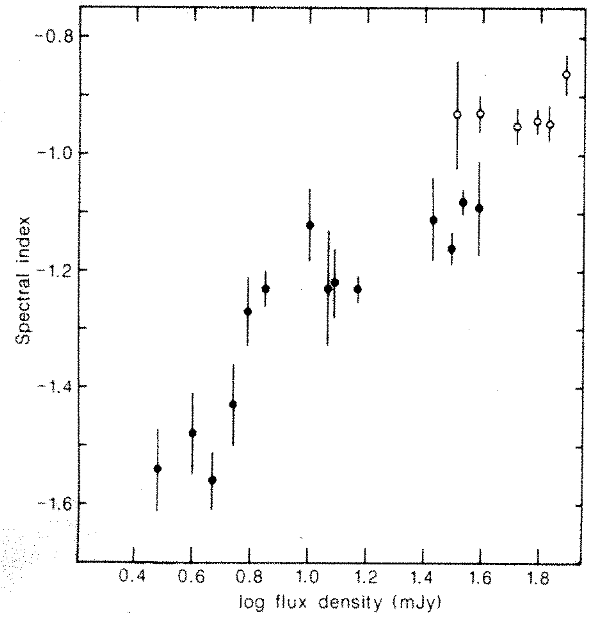


Fig. 2 A plot of infrared spectral index versus the log of the J-band flux showing a very good correlation; as the source gets fainter, the spectrum gets steeper, and vice versa. ●, JHKL spectra indices; ○, data from refs 1, 7, 8 taken at the peak of the 1983 outburst.

process. If the acceleration process produces an electron energy index of ~ -2.7 then an increase in the injection of electrons, or reacceleration of those already present, will result in an increase in emission with a spectral index ~ -0.85, such as was observed during the outburst in 1983 (ref. 1). These electrons will rapidly lose energy radiatively, however, (the results from ref. 6 suggest a magnetic field ~ 1 G in Blazars, corresponding to a lifetime measured in days for electrons radiating at $1\,\mu$m). The effective energy index will then steepen as the higher energy electrons are affected first and the observed spectrum will also steepen as it decays. In support of this conclusion we also note that there is some suggestion of a flattening in the correlation between the spectral index α_{JHKL} and $\log(S_J)$ (the log of the J-band flux) at high flux levels; this is to be expected because at high fluxes we are observing the native electron spectrum, unaffected by energy losses; the observed spectrum cannot become flatter than this.

Hanson and Coe[9] have found a similar correlation for the ultraviolet emission from OJ287 over the period 1983–84. Their ultraviolet spectral indices were generally steeper than those we see in the infrared (except when the infrared flux is very low, when the spectral index steepens to a value comparable with that seen in the ultraviolet). This implies that the ultraviolet spectrum is almost always steepened by energy losses, whereas in the infrared the true injected spectrum can sometimes be observed. We also note that from their optical UBVRI monitoring data in 1983, Moles *et al.*[4] find that the correlation between flux level and spectral index improves as one goes from the infrared into the optical region. This is probably an artefact of the spectral steepening over that wavelength range; the relative variations are larger at shorter wavelengths, due to shorter electron lifetimes, so for a limited data set the correlation will therefore show up more clearly.

The facts that the periods of increased flux after 1983 never reach the values of that outburst and that the spectral index never again reached the value of −0.85, suggest that subsequent events were not as powerful as that which caused the 1983 outburst. Continued infrared monitoring of OJ287 is essential to fully understand the emission from this extremely variable object. In particular, a long campaign of night-to-night monitoring would reveal whether this correlation is maintained on all timescales, or whether the source can take up some additional state. It would be particularly interesting to monitor OJ287 during a very low phase to discover if there was an underlying less variable component such as we have recently discovered[10] in 3C273.

OJ287 is an ideal source for this type of study, being highly variable, relatively bright and a Bl lac object so that one can directly observe the properties of the non-thermal emission over the whole spectrum. However, we note that after very careful removal of the recombination emission, Barr *et al.*[11] found a very similar correlation in the ultraviolet continuum emission from the Seyfert galaxy NGC3783. It is possible that the non-thermal emission from all active nuclei arises by a similar process, but that in many classes of object the behaviour is masked by relatively strong thermal and recombination emission.

We thank the UKIRT staff for their continued assistance with observations, particularly those obtained in service mode. W.K.G. and L.M.J.B. acknowledge financial support from SERC. UKIRT is operated by the Royal Observatory, Edinburgh, on behalf on SERC.

Received 7 July; accepted 3 October 1986.

1. Holmes, P. A. *et al. Mon. Not. R. astr. Soc.* **211,** 497–506 (1984).
2. Corso, G. *et al. Astr. Astrophys.* **140,** L23 (1984).
3. Sillinpaa, A. *et al. Astr. Astrophys.* **147,** 67–70 (1984).
4. Moles, M., Garcia-Pelayo, J. & Masegosa, J. *Mon. not. R. astr. Soc.* **211,** 621–627 (1984).
5. Gear, W. K. *et al. Astrophys. J.* **291,** 511–517 (1985).
6. Gear, W. K. *et al. Astrophys. J.* **304,** 295–204 (1986).
7. Landau, R. *et al. Astrophys. J.* **308,** 78–92 (1986).
8. Roelling, T., Becklin, E. E., Impey, C. D. & Werner, M. W. *Astrophys. J.* **304,** 646–650 (1986).
9. Hanson, C. G. & Coe, M. J. *Mon. Not. R. astr. Soc.* **217,** 831–842 (1985).
10. Robson, E. I. *et al. Nature* **323,** 134–136 (1986).
11. Barr, P., Willis, A. J. & Wilson, R. *Mon. Not. R. astr. Soc.* **202,** 453–466 (1983).

A new low-dimensional metal, $Cs[Pd(S_2C_2(CN)_2)_2]\cdot 0.5\,H_2O$

I. D. Parker*, R. H. Friend*, P. I. Clemenson† & A. E. Underhill†

* Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, UK
† Department of Chemistry and Institute of Molecular and Biomolecular Electronics, University College of North Wales, Bangor LL57 2UW, UK

The discovery of a low-temperature superconducting state in organic compounds of the type $(TMTSF)_2ClO_4$ ($T_c = 1.2$ K) and $(BEDT\text{–}TTF)_2AuI_2$ ($T_c = 4$ K) (where TMTSF is tetramethyltetraselenafulvalene, BEDT–TTF is bis(ethylenedithiolo)tetrathiafalvalene and T_c is the superconducting transition temperature) has stimulated the search for new materials that may show higher values of T_c (refs 1–3). The general problem encountered in molecular charge-transfer salts of this type, which have conduction bands formed by intermolecular overlap of π-electron systems, is that conduction is usually quasi-one-dimensional, with good conduction along the stacking direction. Metals with this one-dimensional character are unstable, and undergo a Peierls transition[4] to a semiconducting state at low temperatures. The relatively few exceptions (mentioned above), which remain metallic down to low temperatures, are considered to do so because they show stronger interstack interactions. We report here a new material with inherently two-dimensional interactions between the molecular π-electron systems and which we are able to stabilize as a metal down to low temperatures (1.4 K) under hydrostatic pressure (12 kbar).

Molecular metals based on square coplanar metal complexes of the KCP ($K_2[Pt(CN)_4]Br_{0.30}\cdot 3H_2O$) type have been known for the past 15 years[5]. More recently the range of compounds with metallic properties has been extended to metal-dithiolene

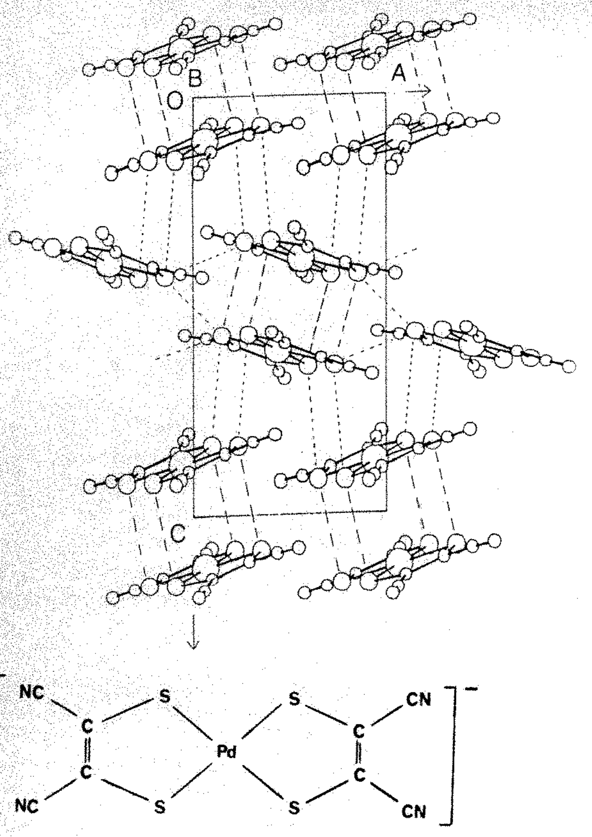


Fig. 1 *b*-Axis projection of the crystal structure of Cs[Pd($S_2C_2(CN)_2)_2$]·0.5 H_2O. The short S-S contacts are shown by dashed lines. Also shown is the molecular structure of [Pd($S_2C_2(CN)_2)_2$]$^-$.

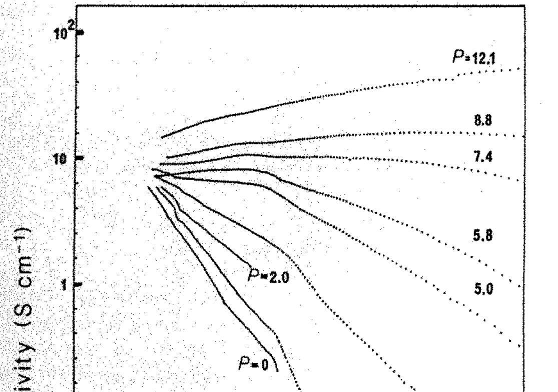


Fig. 2 Conductivity plotted against reciprocal temperature at several pressures (in kbar) for Cs[Pd($S_2C_2(CN)_2)_2$]·0.5 H_2O.

complexes[6]. Figure 1 shows molecular structure of the system of interest here. These are of particular interest because the intermolecular delocalization involves the π-electron systems on the ligands rather than the metal d orbitals (as in KCP). Up to now, all those molecular metals based on metal complexes have been of non-integral oxidation state; the chemical non-stoichiometry of the compound results in partial charge on the complex anion, the conduction band is consequently only partly filled, and metallic behaviour results. For a stoichiometric complex with a single (integral) charge on the anion, the Peierls distortion is of the form of a simple stack dimerization. This distortion is usually very strong and metallic properties are not expected in such materials, even at room temperature. There are several examples of such semiconducting behaviour[7]. Nevertheless, in Cs[Pd($S_2C_2(CN)_2)_2$]·0.5 H_2O we have discovered the first example of an integral-oxidation-state complex with metallic properties.

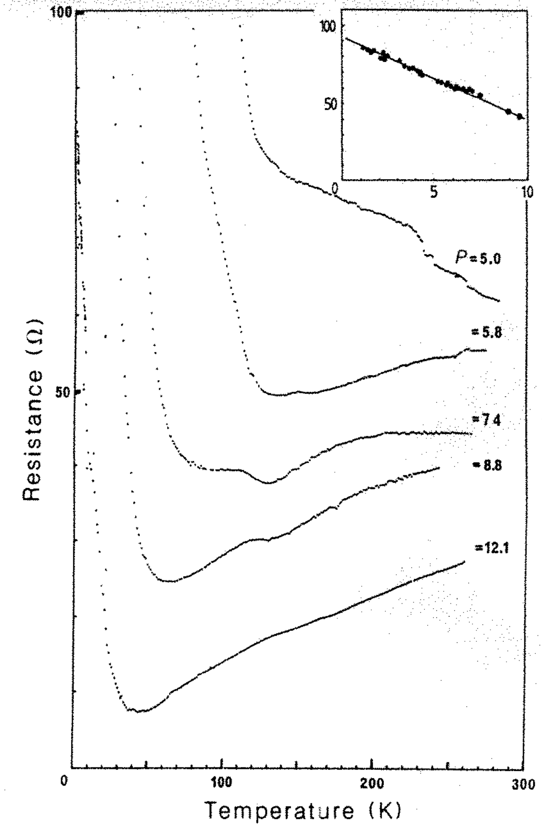


Fig. 3 Resistance plotted against temperature for Cs[Pd($S_2C_2(CN)_2)_2$]·0.5 H_2O. Inset shows the variation of resistance in the region 0–10 K.

Black plate-like crystals of Cs[Pd($S_2C_2(CN)_2)_2$]·0.5 H_2O were prepared by electrocrystallization from an aqueous solution of the dianion salt. Chemical analysis showed the compound to be stoichiometric. The crystal structure has been determined[7]; the structure consists of sheets of the [Pd($S_2C_2(CN)_2)_2$]$^-$ anions separated by layers of Cs^+ and H_2O along the *b* direction. Figure 1 shows the projection of the [Pd($S_2C_2(CN)_2)_2$]$^-$ anions normal to the sheet, and reveals the two-dimensional character of the structure. The [Pd($S_2C_2(CN)_2)_2$]$^-$ anions form dimers in an eclipsed configuration with an intra-dimer distance of 3.329(5) Å, but these are arranged so that there are short S-S contacts between dimers, 3.807 Å along the *c* direction but only 3.576 Å along the *a* direction.

At ambient pressure this material has a room-temperature conductivity of about 5 S cm^{-1} but it is semiconducting at lower temperatures[7]. For measurements under pressure we used a Be-Cu high-pressure cell capable of operation at up to 12 kbar in the temperature range 1.4–300 K (ref. 8). Four terminal measurements of the *a* axis conductivity were made using gold paste contacts. The conductivity at room temperature rises rapidly with pressure, but the changes at lower temperatures are more dramatic, with a transition from semiconducting to metallic behaviour. Results are shown in Figs 2 and 3. Figure 2 plots conductivity against reciprocal temperature, and the

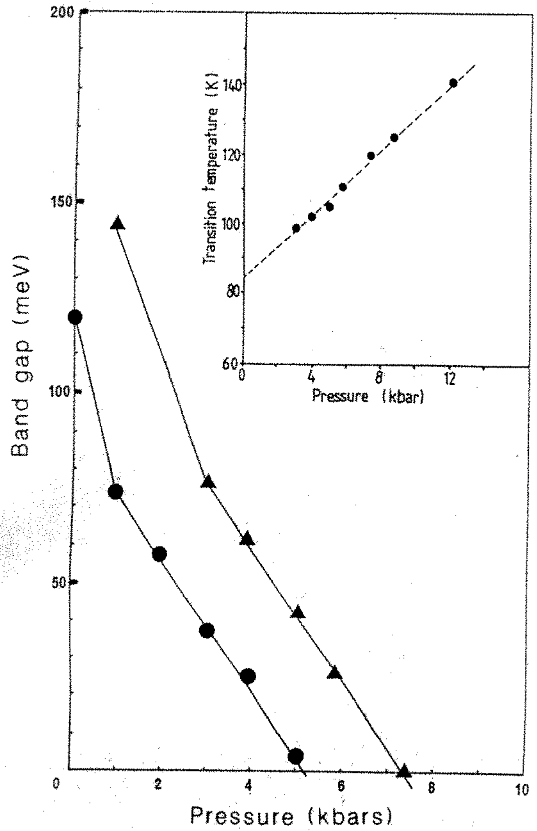


Fig. 4 Activation energies and transition temperature for the semiconducting phases of $Cs[Pd(S_2C_2(CN)_2)_2]\cdot 0.5\ H_2O$ as a function of pressure (band gap = 2 × activation energy). ●, Higher-temperature phase; ▲, lower-temperature phase.

(negative) straight-line slopes give the activation energies for conduction in the semiconducting regime. There is a monotonic decrease in activation energy to zero at about 7 kbar. Figure 3 plots resistance against temperature and shows the linear variation expected for simple metals over much of the temperature regime. At low temperatures the resistance shows an upturn; at 8.8 kbar and below this gives semiconducting behaviour. At the highest pressure investigated, 12.1 kbar, the resistance rise below 40 K is very weak, as shown more clearly in the inset to Fig. 3, and the conductivity at 1.4 K is still above 10 S cm^{-1}.

We have thus established that the conductivity can be kept at metallic levels under pressure but we need to understand why this is not so at ambient pressure. There are clues to this in the behaviour in the semiconducting regime, which is best seen in Fig. 2. There is considerable structure in the data shown, with at least two clearly defined slopes and a sharp break between them. Figure 4 shows the pressure dependence of the activation energies and of the temperature of the transition between them (chosen to be at the peak value of $\partial \ln(\text{conductivity})/\partial(1/T)$). The transitions between these phases are not of the Peierls type. For the Peierls transition there is a linear scaling between the size of the semiconducting gap and the temperature at which the transition occurs[9], whereas we see here that the transition temperature rises while the activation energies fall. We consider instead that the transitions involve some small molecular reorientations which set up appropriate superlattice potentials to 'accidentally' introduce energy gaps in the electron energy bands at the Fermi energy. We think that the effect of pressure may be to increase the intermolecular interactions to such an extent that although the structural transitions may still occur, their effect on the electron levels is diminished. Low-temperature structural measurements are required to follow this.

Received 15 August; accepted 16 October 1986.

1. Int. Conf. Synth. Metals, Kyoto, June 1986 *Synth. Metals* (in the press).
2. Williams, J. M. *Prog. inorg. Chem.* **33**, 183–220 (1985).
3. Wudl, F. *Acct. chem. Res.* **17**, 227–232 (1984).
4. Peierls, R. E. *Quantum Theory of Solids* 108–114 (Oxford University Press, London, 1955).
5. Underhill, A. E., Williams, J. M., Schultz, A. J. & Carneiro, K. in *Extended Linear Chain Compounds* Vol.1 (ed. Miller, J. S.) 73–118 (Plenum, New York, 1982).
6. Ahmad, M. M. *et al. Phys. Rev.* **B29**, 4796–4799 (1984).
7. Clemenson, P. I. *et al.* in Int. Conf. Synth. Metals, Kyoto, June 1986 *Synth. Metals* (in the press).
8. Guy, D. R. P. & Friend, R. H. *J. Phys.* E. **19**, 430–433 (1986).
9. Friend, R. H. & Jerome, D. *J. Phys.* C **12**, 1441–1477 (1980).

Organ pipe radiant modes of periodic micromachined silicon surfaces

Peter J. Hesketh*, Jay N. Zemel* & Benjamin Gebhart†

* Department of Electrical Engineering and † Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6390, USA

In recent studies on small pyroelectric thermal anemometers with roughened surfaces we showed that one of the most widely used heat transfer models[1,2] yielded calculated anemometer responses for flow and geometric behaviour that agreed functionally with observations, but were significantly smaller than the experimental data[3–5]. As the first stage in investigating the role of small structures in heat transfer, we initiated a study of emittance from deep gratings. Here we report measurements at 400 °C of infrared ($3\ \mu m \leqslant \lambda \leqslant 14\ \mu m$), normal, s- and p-polarized spectral emittances of 45 μm deep, near square-wave gratings of heavily phosphorus doped (110) silicon (P content ~5×10^{19} cm^{-3}). The grating surface repeat scales, Λ, were 10, 14, 18 and 22 μm, yielding a range of Λ/λ from 0.14 to 7.33. The s-polarization vector was parallel to the grating slots. Both s and p spectral emittances had pronounced resonant periodicities with a characteristic length of ~42 μm. A reasonable explanation for this behaviour is the presence of standing waves in the air slots perpendicular to the silicon surface similar to those in an organ pipe. While the resonant amplitude of the s polarization does not depend significantly on Λ it does for the p polarization. No explanation for the Λ dependence of the p polarization is known.

Although radiant emission from macroscopic black bodies and flat, smooth metallic surfaces is well understood[6,7], emission from deep, near square-wave gratings when Λ is comparable to or <λ is not. The theory of radiant emission from structures whose scale, S, is comparable to or <λ, has drawn considerable attention[8–18]. Measurements of semiconductors emissivity gave information on the optical properties of bulk dopants[19–21]. The effect of surface asperities on metal emissivity has been thoroughly studied[22–27]. Generally, the polarized emissivity of rough metallic surfaces depends on the r.m.s. value of surface asperities with the energy concentrated at small polar angles[24–28]. The work on both s- and p-polarized photo-acoustic spectrum from a silvered bigrating is the closest to our study, but the depth (H) used was still ≪Λ (ref. 29 and see ref. 30).

Our experiments employed (110) oriented silicon wafers on which square-wave gratings were etched to a common depth of ~45 μm using standard photolithography and a 40% (w/v) solution of KOH at 52 °C. After etching, the microconfigured wafers were spin-coated with a phosphorus based dopant and heated at 1,250 °C for 10 h. The peak phosphorus surface concentration is in the $0.4\text{–}0.6\times10^{20}$ cm^{-3} regime from sheet resistance measurements. This carrier concentration yields an effective skin depth of ≤4 μm, for λ > 5 μm. Auger measurements gave surface phosphorus concentrations of ~1×10^{20} cm^{-3}, in good agreement with the sheet resistance measurements. The heavily doped back surfaces of the silicon wafers were used as heater elements. Figure 1*a* shows the grating and the polarizations studied. Figure 1*b* shows a scanning electron micrograph of a

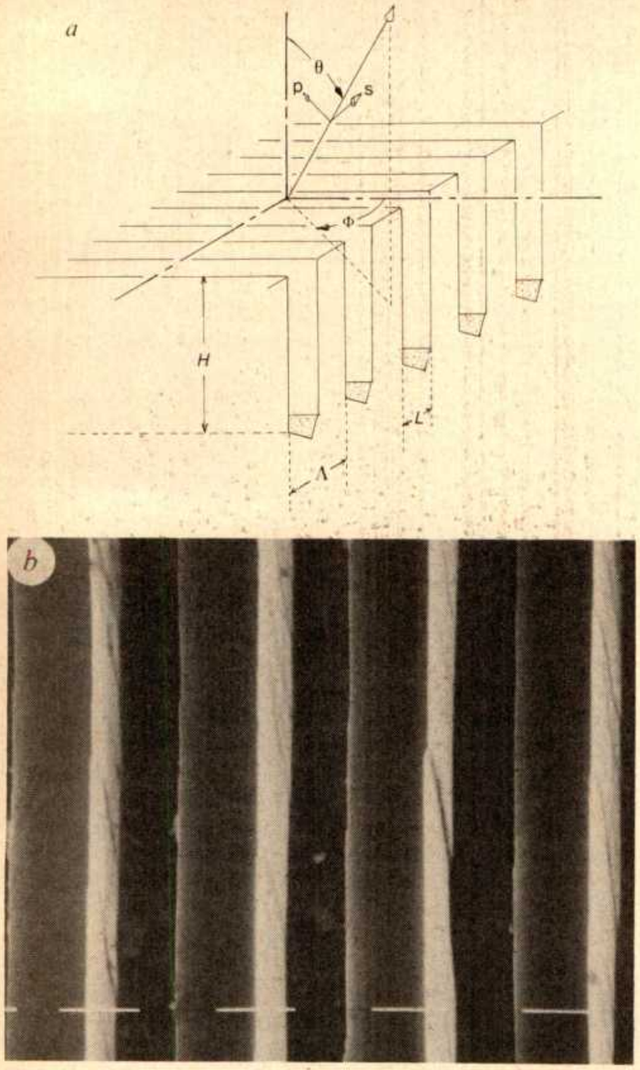


Fig. 1 *a*, Schematic drawing of the grating. Λ, repeat distance; *L*, slot width; *H*, depth; s- and p-polarization vectors; *θ*, polar angle and Φ, azimuthal angle. *b*, Electron micrograph of the microgrooves; $H=\Lambda=22\ \mu\text{m}$, at a magnification of 1,000×. A 10-μm-wide marker is shown in white.

grating with $\Lambda=22\ \mu\text{m}$. Chips (1×1 cm) containing the microconfigured surface were mounted in a specially designed low-thermal-loss holder. More energy was required to reach operating temperatures for microconfigured surfaced samples compared to the smooth surfaces. A black-body source was constructed from a copper cylinder. The temperature of the source and the temperature difference between it and the sample were monitored throughout the measurement sequence. The 400 °C black-body source, controlled to better than 0.01 °C, was used as a reference for stabilizing the silicon sample temperature. Temperature differences between the silicon sample and the black body were <0.5 °C. Temperature fluctuations in the sample were typically <0.1 °C during a measurement. The emissivity measurement system is shown in Fig. 2. The signal-to-noise ratios of the infrared detector were ~100:1 or better at the higher temperatures. Data obtained at lower temperatures agreed well with these observations. The data were processed with a small data logging system based on a single board Z-80 computer.

The measurement procedure consisted of logging the polarized spectral intensity of the 400 °C black-body and silicon grating, both at normal incidence. This corresponded to a polar angle, $\theta=0°$, and an azimuthal angle, $\Phi=90°$. The polarized

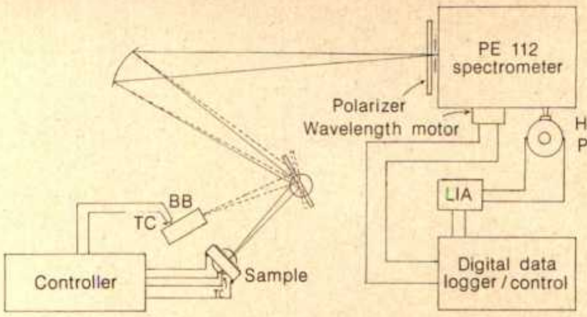


Fig. 2 Schematic layout of the emittance measuring equipment: TC, thermocouple; BB, black-body; LIA, lock-in amplifier; PE, Perkin Elmer.

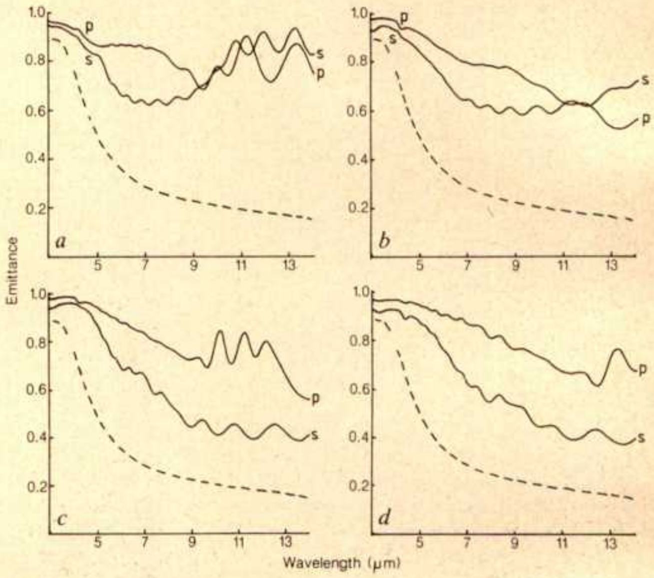


Fig. 3 Normal polarized, $T=400\,°\text{C}$, spectral emittance as a function of spectral wavelength for the p- and s-polarizations. For comparison, the smooth surface emittance is shown as the dashed line. All gratings had a depth $H=45\pm2\ \mu\text{m}$: *a*, $\Lambda=10\ \mu\text{m}$; $L=7.3\ \mu\text{m}$; *b*, $\Lambda=14\ \mu\text{m}$, $L=8.4\ \mu\text{m}$; *c*, $\Lambda=18\ \mu\text{m}$; $L=12.6\ \mu\text{m}$; *d*, $\Lambda=22\ \mu\text{m}$, $L=14\ \mu\text{m}$.

spectral emittance, $\varepsilon(^{\mathrm{s}}_{\mathrm{p}};\lambda;T=400\,°\mathrm{C})$ is defined as the ratio of these spectral intensities. For the above angles, the s-polarization vector is parallel and the p-polarization is perpendicular to the slots (Fig. 1*a*). Shifting Φ to 0° reverses the definitions of the s- and p-polarization vectors relative to the slots so they are perpendicular and parallel to the slots, respectively. The reproducibility of ε was 2–4% between runs. Pronounced periodic oscillations were found in both $\varepsilon(\mathrm{p};\lambda;400\,°\mathrm{C})$ and $\varepsilon(\mathrm{s};\lambda;400\,°\mathrm{C})$ from gratings with $\Lambda=10\ \mu\text{m}$, 14 μm, 18 μm, and 22 μm (Fig. 3). The value of ε for a smooth, heavily doped, silicon surface is shown in Fig. 3 as a dashed curve. Temperature had little influence on the magnitude of ε and essentially none on the period or amplitude of the observed oscillations.

The wavenumber, $K^{(i)}$, corresponding to maxima in the polarized spectral emittance ($K_{\mathrm{m}}^{(i)}=1/\lambda_{\mathrm{m}}^{(i)}$, $i=\mathrm{s},\mathrm{p}$) were plotted against an integral mode number, $m^{(i)}$ (for the appropriate polarization, $i=s,p$). The origin of the $K_m^{(i)}$ plot for the individual gratings was adjusted so that the data overlapped as shown in Fig. 4. Note that there is essentially no influence of Λ on the slope. We constructed a model in which the $K_{\mathrm{m}}^{(i)}$ are related to the $m^{(i)}$ by

$$K_{\mathrm{m}}^{(i)}=\begin{cases} m^{(i)}/2H \\ (2m^{(i)}+1)/4H \end{cases} \qquad (1)$$

where the upper relation applies to modes where the nodes

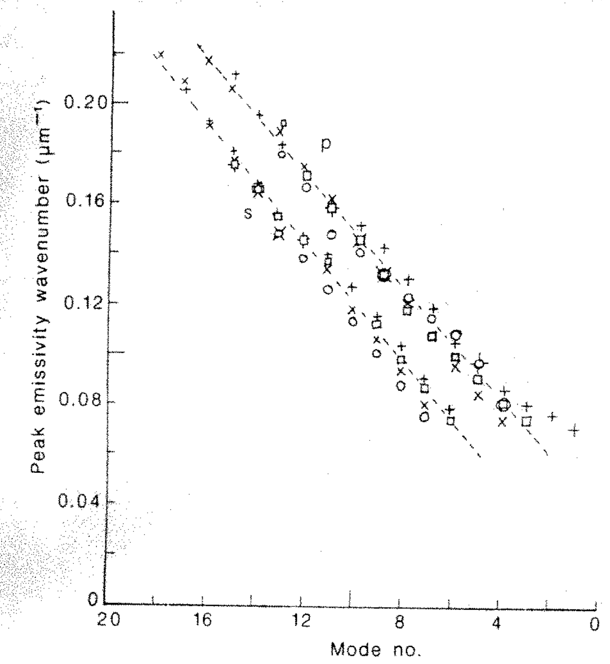


Fig. 4 Peak emittance wavenumber versus mode number; p- and s-polarizations; $H = 45 \pm 2$ µm; Λ, ○ = 10 µm; ⊙ = 14 µm; + = 18 µm; × = 22 µm.

occur at 0 and H and the second where one node occurs at 0 and the anti-node is at H. In both cases,

$$\Delta K_{\mathrm{m}}^{(i)} = \Delta m^{(i)}/2H \qquad (2)$$

From Fig. 4, $H = 42$ µm in striking agreement with the average depth of ~45 µm measured for the four different gratings. This is even more remarkable because Λ varies by a factor of >2. Whereas the slope of $K_{\mathrm{m}}^{(i)}(m^{(i)})$ is independent of Λ, inspection of the amplitudes of the emittance oscillation of Fig. 3 suggests that the p-polarized data oscillations are far more sensitive to Λ than are the s-polarized. We do not yet have any explanation for this difference.

The s-polarized emittance oscillations have no parallel in either experiment or theory. Work by Inagaki *et al.* suggests that periodic resonant absorptance occurs on a bigrating, but their data indicate that the magnitude of the absorptance is much smaller for the s-polarization than for the p-polarization[29]. Recent calculations show that substantial field enhancement arises in the slots of a deep cavity for the s-polarizations[14,18]. However, the far-field pattern associated with these s-polarized radiative modes is weak and almost identical to those from a smooth mirror of similar composition. In our structures, thermal energy populating the radiative (or absorptive) modes of the deep grating structure may be coupled out by asperities in the slot geometries. In any case, the peaks in the emittance are due to emission from 'organ pipe' or planckian type resonant modes in the slots of the grating.

We thank Judson Infrared for the HgCdTe infrared detector, Dr A. Ipri for supplying the photomasks; the EEMS Program of the NSF for partial support, D. Silversmith, Program Manager, under grant ECS-84-12241 (P.H., J.N.Z.); partial support from NSF grant CBT-84-18517 (B.G.); and Professors H. Baltes and A. Maradudin for useful discussions.

Received 10 June; accepted 10 September 1986.

1. Schlichting, H. *Boundary Layer Theory* (McGraw Hill, New York, 1968).
2. Gebhart, B. *Heat Transfer* (McGraw Hill, New York, 1961).
3. Rahnamai, H. & Zemel, J. N. *Sensor Actuat.* **2,** 1–16 (1981).
4. Frederick, J. R., Zemel, J. N. & Goldfine, N. F. *J. appl. Phys.* **57,** 4936–4943 (1985).
5. Hesketh, P., Zemel, J. N. & Gebhart, B. *J. appl. Phys.* **57,** 4944–4949 (1985).
6. Planck, M. *Theory of Heat Radiation* (Eng. trans. of *Theorie der Warmestrahlung*, 2nd edn 1913) (Dover, New York, 1959).
7. Siegel, R. & Howell, J. R. *Thermal Radiation Heat Transfer* 2nd edn (McGraw Hill, New York, 1981).
8. Baltes, H. P. *Infrared Phys.* **16,** 1–8 (1976).
9. Derrick, G. H., McPhedran, R. C., Maystre, D. & Neviere, M. *Appl. Phys.* **18,** 39–52 (1979).
10. Elson, J. M. & Sung, C. C. *Appl. Opt.* **21,** 1496–1501 (1982).
11. Chang, R. K. & Furak, T. E. (eds) *Surface Enhanced Raman Scattering* (Plenum, New York, 1982).
12. Fano, U. *J. opt. Soc. Am.* **31,** 213–222 (1941).
13. Hessel, A. & Oliner, A. A. *Appl. Opt.* **4,** 1275–1297 (1965).
14. Wirgin, A. & Maradudin, A. A. *Phys. Rev.* **B31,** 5573–5576 (1985).
15. Maystre, D. in *Electromagnetic Surface Modes* (ed. Boardman, A. D.) 166 (Wiley, New York, 1982).
16. Glass, N. E., Maradudin, A. A. & Celli, V. *Phys. Rev.* **B27,** 5150–5153 (1983).
17. Garcia, N. & Maradudin, A. A. *Opt. Commun.* **45,** 306–310 (1983).
18. Wirgin, A. & Maradudin, A. A. *Appl. Surf. Sci.* (submitted).
19. Stierwalt, D. L. & Potter, R. F. *Proc. int. Conf. on the Physics of Semiconductors, Exeter*, 513 (Institute of Physics and the Physical Society, London, 1962).
20. Stierwalt, D. L. *Appl. Opt.* **5,** 1911–1915 (1966).
21. Stierwalt, D. L. & Potter, R. F. in *Semiconductors and Semimetals* Vol. 3 (eds Willardson, R. K. & Beer, A. C.) (Academic, New York, 1968).
22. Lofving, S. *Appl. phys. Lett.* **36,** 632–633 (1980).
23. Edwards, D. K. *Advances in Thermophysical Properties at Extreme Temperatures and Pressure, 3 int. Symp.* 189 (ASME, New York, 1965).
24. Kanayama, K. *Heat Transfer—Jap. Res.* **1,** 11–21 (1972).
25. Rolling, R. E. *Progress in Astronautics and Aeronautics: Thermophysical Properties of Spacecraft and Planetary Bodies* (ed. Heller, G. B.) (Academic, New York, 1967).
26. Craig, S. & Harding, G. L. *Sol. Energy Mater.* **4,** 245–264 (1981).
27. Petit, G. D., Cuomo, J. J., DiStefano, T. H. & Woodall, J. M. *IBM J. Devl.* **22,** 372–377 (1978).
28. Church, E. L., Jenkinson, H. A. & Zawada, J. M. *Opt. Engng.* **16,** 380–385 (1977).
29. Inagaki, T. Goudennet, J. P., Little, J. W. & Arakawa, E. T. *J. opt. Soc. Am.* **B2,** 433–439 (1985).
30. Baltes, H. P. & Hilf, E. R. *Spectra of Finite Systems* (Bibliographisces Institute, Zurich, 1976).

Evidence for a link between atmospheric thermonuclear detonations and nitric acid

G. Holdsworth

National Hydrology Research Institute and
Arctic Institute of North America, University of Calgary,
2500 University Drive N.W., Calgary, Alberta, Canada T2N 1N4

Suitably located glacier cores, obtained from high-altitude, low-temperature sites, can reveal detailed information about atmospheric air chemistry at sub-annual resolution[1]. Such data may provide input to climate-change models, the study of acid precipitation patterns and many other phenomena. Here I present data from an ice core which show that during the era of intense atmospheric thermonuclear weapons testing (ATWT) a significant part of the nitrate content in the snow was modulated by the intensity of the nuclear detonations. The fixation of nitrogen by nuclear fireballs leads to NO_x gases in the atmosphere[2] and ultimately to nitric acid in precipitation. At certain concentrations, these gases and the associated aerosols may perturb the climate[3,4].

In 1980 a series of snow and ice cores were retrieved from an altitude of 5340 m (approximately the 500 mbar or mid-troposphere level) on the upper plateau of Mt Logan (60° 36′ N; 140° 30′ W) in the Yukon Territory, Canada. The recorded mean annual snow temperature is −28.5 °C and the mean net annual accumulation rate is ~0.37 m water equivalent. The gross β-activity levels in the upper part of one of these cores were determined to a depth of almost 27 m, corresponding to about the year 1943 (ref. 1). These data were interpreted in terms of the known chronology of ATWTs[1,5].

Analyses for the major ions in the firn were performed at the Laboratoire de Glaciologie, CNRS, Grenoble, for a section of core between 10 and 21 m (AD 1967–1953)[6]. Inspection of the nitrate concentration data and the gross β-activity curve showed that these two data sets shared several major features: a prominent 1963 maximum preceded by a ~1960 (temporary moratorium) minimum and followed by another (post-1962 limited test ban treaty) minimum in ~1966. Analyses of the upper 10 m of core, with overlap on the lower section, were then carried out at the University of Calgary, using the same precautions exercised earlier[6,7]. For these purposes a parallel

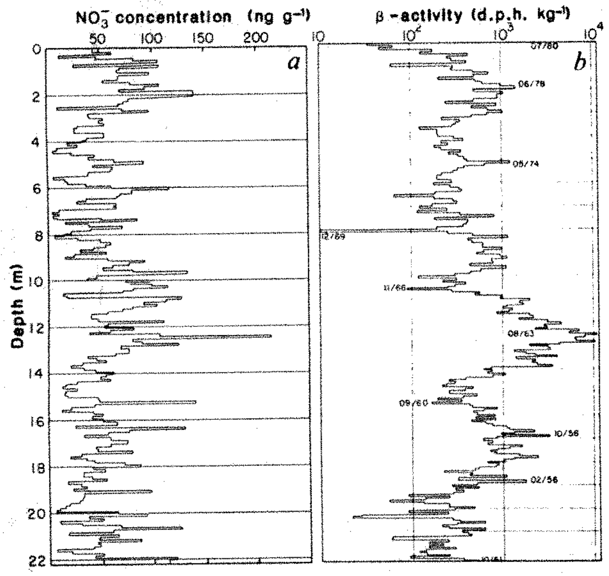


Fig. 1 *a*, Raw data for nitrate concentration and *b*, gross β-activity levels plotted against depth in the core. The top of the core corresponds to mid-1980 and the bottom to ~1952. Prominent peaks and troughs in *b* have dates assigned according to ref. 1.

core taken within 1 m of the earlier core[6] was analysed, thus providing a check on the consistency of results. At the 0.10 m cutting interval used, spikes in the nitrate and other anion data are quite common and cannot be associated with contamination. Contamination is further discussed in the section dealing with the origin of nitrates in the core. A comparison was then made between the upper 10 m of nitrate data and the gross β-activity curve (Fig. 1). This comparison revealed that, at least from 1973 to 1980, the People's Republic of China ATWTs were tending to modulate the NO_3^- levels.

Because the Chinese and the preceding Soviet ATWTs were done at locations with the same general latitude (±15°) as Mt Logan, and because the latter represents a relatively near-field site, it is not unexpected that the chemistry of the core should be influenced by these ATWT events. It has already been shown[8] that ATWT events produce copious quantities of nitric oxide (NO) (~10^{10} g Mt^{-1}). Although most of the detonations were made in the troposphere, large quantities of NO were injected, with other explosion products, into the stratosphere, where the residence time of the gases and aerosols could be up to several years[8].

A timescale for the core was established[1] to an accuracy of between ±0.25 yr and ±0.5 yr. The original data (typically 10 samples per year) were then converted into 6-monthly mean values to produce the time series of gross β-activity and $[NO_3^-]$ shown in Fig. 2. This particular time interval used for averaging represents the shortest unit that can be justifiably used while still preserving significant structure in the time series. Further smoothing was carried out for cross-correlation purposes. An attempt was made to reconstruct the gross β-activity values to their original values at the time the snow was deposited. This would allow a closer comparison to be made between gross β-activity and $[NO_3^-]$ with respect to amplitude information. The dashed blocks in Fig. 2*b* correspond to a simple radioactive decay correction based on the assumption that the dominant radionuclides present were ^{90}Sr and ^{137}Cs. This assumption is evidently not a good approximation for the most recent 8 yr of data[9], but reconstructions made using a sample-age-dependent effective half life[9] yield results at variance with the gross β-activity data[1] for air. Errors introduced by this adjustment of the data will not materially affect the major conclusions of this paper.

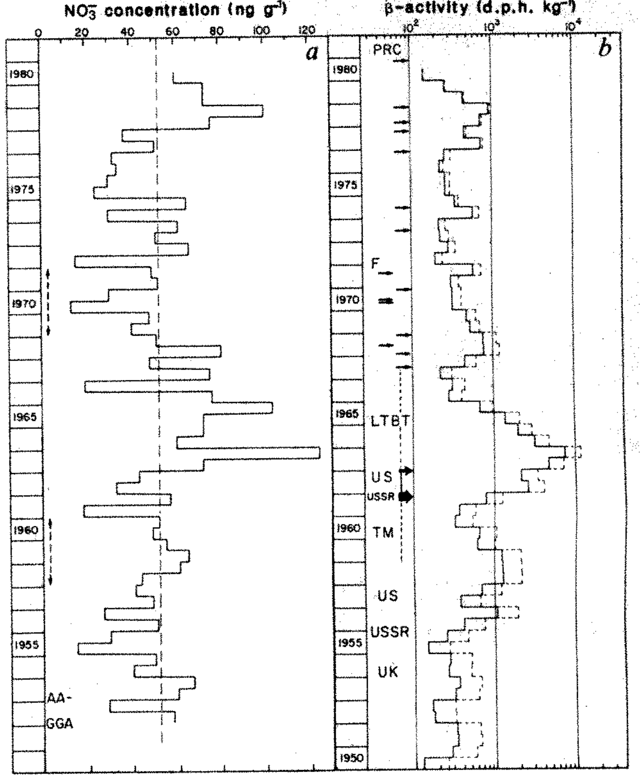


Fig. 2 Time-series plots of *a*, nitrate concentration, where broken arrows indicate position of auroral activity (AA) or global geomagnetic activity (GGA) maxima[20] and *b*, gross β-activity levels, where the arrows indicate the dates of prominent atmospheric thermonuclear weapons testing (PRC, People's Republic of China; F, France; US, United States; USSR, Union of the Soviet Socialist Republic; UK, United Kingdom; TM, Temporary Moratorium; LTBT, Limited Test Ban Treaty). The dashed blocks above the solid blocks indicates an approximate reconstruction of the original gross β-activity levels at the time of deposition.

Cross-spectral analyses were performed on the two time series. From the measured values of the gross β-activity, *A*, the value of the cross-correlation coefficient, *r*, between the two series is +0.42 at zero lag and +0.45 to +0.48 at 2-2.5 yr lag. From the decay-corrected gross β-activity values, A_0 the corresponding values of *r* are +0.41 and +0.44 to +0.47 respectively, all significant at the 95% level. For other lags the correlation is not significant. The coherence at zero lag was expected, as the two time series exhibit this feature visually. A clear explanation for the apparent correlation at 2-2.5 yr lag cannot be offered. It may represent a statistical artefact.

Several studies[10-13] concerned with the origin of nitrate in snow, firn and ice cores give twelve possible sources of this major trace ion, including atmospheric thermonuclear explosions. Local contamination of the core is not considered applicable here. Strict procedures and controls employed during sampling and analysis ensured that nitrate contamination was not significant. Measurements were made at two independent research institutions and a third commercial laboratory as a check by two different methods (ion chromatography and colorimetry). The core was independently sampled for gross β-activity, and as that curve resembles quite closely other published curves from different regions of the world[1,5], it is concluded that contamination cannot be significant. Finally, variations in nitrate in the core may also result from changes in accumulation rate[12], ablation, or meteorological parameters not associated with these quantities.

It is not possible to quantify each of the above sources in a detailed manner and it is furthermore evident that relative contributions from these sources vary geographically, both regionally[7,10-12] and with altitude[13,14]. It was shown[7] that no detectable trends in either $[SO_4^{2-}]$ or $[NO_3^-]$ exist in the Mt Logan snow from ~1880 to 1980, thus evidently eliminating anthropogenic input related to fuel combustion, as interpreted from ice cores from lower-altitude Greenland and the NE Canadian Arctic[15-17]. Therefore the most likely sources for significant nitrate input to Mt Logan snow in the specific time interval considered here are: (1) ionization in the upper atmosphere from various sources[10,11], (2) meteoroid fixation[10,11], (3) photochemical fixation[10,11,18], (4) volcanic emissions[7,19] and (5) atmospheric thermonuclear weapons testing[8]. The one mechanism for producing ionization in the upper atmosphere for which there is an established cyclicity (sunspot-auroral activity) does not appear to produce a signal in these data or in the Greenland data[12]. In Fig. 2*a* the broken arrows indicate periods of maximum auroral and geomagnetic activity[20]. No data are available for evaluating sources (2) and (3). In addition, the period AD 1952-80 was not particularly marked by major volcanic eruptions close enough to Mt Logan to cause a significant $[NO_3^-]$ pulse, unlike in 1912 for Novarupta[7]. Most of the relatively large eruptions that occurred in the north Pacific during this period took place on the Kamchatka Peninsula[19]. There is no consistent correlation between these events and the peaks in $[NO_3^-]$ except for Kliuchevskoi (1965, 66) which has a volcanic explosive index (VEI) of 3, whereas others during this period had VEI values of 4 and 5 (ref. 19). However, gaseous emissions are not necessarily related to VEI[21,22]. There were other volcanic events[19] that occurred during the 1964-66 interval that could have contributed cumulatively to the significant peak in $[NO_3^-]$ for AD 1965±0.5, which is particularly anomalous with respect to the ATWT history. Of course, any input of nitrate from these volcanic sources depends quite crucially on air-mass trajectories. A sulphate peak occurs with the nitrate peak for 1965 (ref. 6) so at least some volcanic input is likely.

The generation of NO from ATWT explosions has been put on a semi-quantitative basis[8], thus providing an opportunity for checking whether the core data are consistent with theoretical estimates. The 1963-65 'excess' nitrate is defined as that amount above the mean for 1952-80, or 52 ng g^{-1} (Fig. 2*a*). It will be assumed initially that the 1965-66 spike, which does not correlate with an enhanced gross β-activity level, is due to ATWTs: in 1965 and 1966 the People's Republic of China conducted small ATWTs. This assumption would imply different transport processes or routes for the two atmospheric species being considered. Next, the excess nitrate is converted into a load value by using the net annual snow accumulation data for these years. This calculation yields a total value of 4.45×10^4 µg m^{-2} of NO_3^-, or roughly 4.5×10^4 µg m^{-2} of HNO_3. If this load density is assumed to apply uniformly over the Northern Hemisphere (2.54×10^{14} m^2) then a total of about 11 Tg of HNO_3 is obtained. This is surely an upper limit, because if most of this HNO_3 is derived from ATWT events originating in the USSR, then the load density at the Mt Logan site must be much higher than the average for the Northern Hemisphere, as it is for the gross β-activity[1]. Because it was found that the gross β-activity at Mt Logan was about 32% higher than in Greenland[1], it is reasonably assumed that the Mt Logan nitrate flux from ATWTs is also correspondingly higher than in Greenland. If the Mt Logan nitrate flux derived from ATWTs is assumed to be twice the northern hemispheric average, then the total load for the stated period would be reduced to about 6 Tg. If, now, the excess nitrate during 1965 and up to 1966 is due to local volcanic emissions, then by trimming off the 1965-66 spike to the long-term mean level, a revised total computed load would be close to 5 Tg.

Assuming that a 1 megatonne bomb yield produces ~10^{10} g NO, equivalent to ~2×10^{10} g HNO_3, and that 340 megatonnes of cumulative yield occurred between October 1961 and December 1962 (ref. 8), then the total yield of HNO_3 for this period would be ~6 Tg. Another estimate[2] of cumulative bomb yield for this period is <340 Mt, so that a lower nitrate load is possible. From these rough calculations it may be seen that the excess nitrate for 1963-65, as derived from Fig. 2*a*, is of the right order for it to be accounted for by the known ATWT events that occurred during the preceding period.

In conclusion, the data presented here suggest that the ATWT events of more than two decades ago generated significant quantities of nitric oxide, the precursor gas for the nitric acid found in the snow on Mt Logan. The all-time peak pulse in gross β-activity coincides with the largest nitrate pulse for the period 1952-80. The observed excess nitrate, above the three decades mean for the 1963-65 period, can roughly be accounted for if the theoretical predictions for ATWT-produced NO are correct. These results, together with the observation that the gross β-activity curve is similar to other curves covering the same period, from different parts of both hemispheres[5], strongly suggest a common physical connection between nitrate and the radionuclides responsible for the β-activity in the snow, rather than just a connection through a common transport process. This latter mechanism would almost certainly explain the significant coherence between $[SO_4^{2-}]$ and $[NO_3^-]$ found particularly in the lower 10 m section of the core studied. For these two species $r=+0.49$ at zero lag, significant at the 95% level. Because this applies during the interval that includes the heaviest ATWT detonations, it would appear that the radioactive particles were playing a dominant role in acting as the nuclei for the aerosols that carried both the H_2SO_4 and HNO_3, in accordance with stratospheric aerosol theory[4]. A possible mechanism for injecting the stratospheric aerosols into the troposphere could be tropopause folding[23], a process that may be enhanced by lower-stratosphere overloading. Lastly, more direct evidence suggesting a connection between the 1976 People's Republic of China ATWT to a pulse of atmospheric nitrate in filtered air samples has been reported (W. A. Sedlacek, E. J. Mroz, A. L. Lazrus and B. W. Gandrud, in preparation).

I thank P. Upton and A. Williams, and others associated with the Arctic Institute of North America; the ice core drilling team who made the recovery of the core possible; R. Delmas and W. Sedlacek for helpful discussions; B. Gosling and B. Smith for the statistical processing of the data; and R. VanEverdingen and J. Winchester for significant improvements to the manuscript, a preliminary version of which was presented at the 1985 AGU fall meeting in San Francisco. This paper is published with the approval of the Director, National Hydrology Research Institute, Canada.

Received 7 April; accepted 8 October 1986.

1. Holdsworth, G., Pourchet, M., Prantl, F. A. & Meyerhof, D. P. *Atmos. Envir.* **18,** 461-466 (1984).
2. Carter, M. W. & Moghissi, A. A. *Hlth Phys.* **33,** 55-71 (1977).
3. Turco, R. P., Whitten, R. C. & Toon, O. B. *Rev. Geophys. Space Phys.* **20,** 233-279 (1982).
4. Prospero, J. M. *et al. Rev. Geophys. Space Phys.* **21,** 1607-1629 (1983).
5. Pourchet, M., Pinglot, F. & Lorius, C. *J. geophys. Res.* **88,** 6013-6020 (1983).
6. Delmas, R., Legrand, M. & Holdsworth, G. *Ann. Glaciol.* **7,** 213 (1985).
7. Holdsworth, G. & Peake, E. *Ann. Glaciol.* **7,** 153-160 (1985).
8. Foley, H. M. & Ruderman, M. A. *J. geophys. Res.* **7,** 4441-4450 (1973).
9. Ambach, W. & Rehwald, W. *Arct. alpine Res.* **14,** 163-166 (1982).
10. Parker, B. C., Zeller, E. J. & Gow, A. J. *Ann. Glaciol.* **3,** 243-246 (1982).
11. Laird, C. M., Zeller, E. J., Armstrong, T. P. & Parker, B. C. *Geophys. Res. Lett.* **9,** 1195-1198 (1982).
12. Herron, M. M. *J. geophys. Res.* **87,** 3052-3060 (1982).
13. Lyons, W. B. & Mayewski, P. A. *Geophys. Res. Lett.* **10,** 1160-1163 (1983).
14. Bollinger, M. J. *et al. J. geophys. Res.* **89,** 9623-9631 (1984).
15. Koerner, R. M. & Fisher, D. A. *Nature* **295,** 137-140 (1982).
16. Neftel, A., Beer, J., Oeschger, H., Zurcher, F. & Finkel, R. C. *Nature* **314,** 611-613 (1985).
17. Mayewski, P. A. *et al. Science* **232,** 975-977 (1986).
18. Legrand, M. R. & Delmas, R. J. *Atmos. Envir.* **18,** 1867-1874 (1984).
19. Simkin, T. *et al. Volcanoes of the World* (Hutchinson Ross, Stroudsburg, Pennsylvania, 1981).
20. Legrand, J. P. & Simon, P. A. *Solar Phys.* **70,** 173-195 (1981).
21. Sedlacek, W. A., Mroz, E. J., Lazrus, A. L. & Gandrud, B. W. *J. geophys. Res.* **88,** 3741-3776 (1983).
22. Berresheim, H. & Jaeschke, W. *J. geophys. Res.* 3732-3740 (1983).
23. Shapiro, M. A. *J. atmos. Sci.* **37,** 994-1004 (1980).

Weakening of rock salt by water during long-term creep

Janos L. Urai, Christopher J. Spiers, Hendrik J. Zwart & Gordon S. Lister*

Institute of Earth Sciences, University of Utrecht, Budapestlaan 4,
3508 TA Utrecht, The Netherlands

The rheological properties of rock salt are of fundamental importance in predicting the long-term evolution of salt-based radioactive waste repositories and strategic storage caverns, and in modelling the formation of salt diapirs and associated oil traps[1,2]. The short-term, high-stress rheology of rock salt is well known from laboratory experiments; however, extrapolation to appropriately low stresses fails to predict the rapid flow seen in certain natural structures. Furthermore, experiments have failed to reproduce the recrystallized microstructure of naturally deformed salt. Here we report experiments indicating that the above discrepancies can be explained by taking into account the influence of trace amounts of brine. Trace brine is always present in natural salt but sometimes escapes during experiments. Our tests on dry dilated salt show more or less conventional dislocation creep behaviour, but brine-bearing samples show marked weakening at low strain rates. This is associated with dynamic recrystallization and a change of deformation mechanism to solution transfer creep. Because natural rock salt always contains some brine, these results cast substantial doubt on the validity of presently accepted dislocation creep laws for predicting the long-term rheological behaviour of salt in nature.

Creep of rock salt has been the subject of extensive investigations over the past two decades[1]. Existing data[3,4] show good agreement for a wide range of rock salts, and satisfactorily describe short-term mine convergence[5]. Long-term rheological properties, however, are not accessible in the laboratory, and predictions are generally based on extrapolation. All previous experimental work indicates that 'dislocation creep' is the most important deformation mechanism in salt. This mechanism also seems common during natural deformation, as evidenced by preferred lattice orientations[1,6] and the presence of dislocation substructures[7]. However, many naturally deformed salts also show evidence of (dynamic) recrystallization by grain-boundary migration[8,9], a process rarely observed in experiments[6,10] and generally held to be impossible under natural conditions[1,11,12]. This discrepancy between the microstructure of naturally and experimentally deformed salt casts doubt on the applicability of laboratory flow data to natural deformation, as it suggests a change in microscale processes at low strain rates.

The flow rates of salt glaciers increase dramatically after a few centimetres of rainfall[13,14], to rates several orders of magnitude faster than suggested by experimental data, and the salt apparently undergoes extensive recrystallization. On this basis, it has been proposed that small amounts of brine can dramatically alter the deformation mechanisms and rheology of rock salt[13,15]. As all natural rock salts contain small amounts of brine (usually 0.1-1.0 wt.%)[16], the weakening effects of water operating in glacier salt should also be important in other salt deposits[13,14].

On theoretical grounds[8], even very small amounts of brine can be expected to influence salt creep. First, dynamic recrystallization can be strongly enhanced by thin brine films at grain boundaries[8,17]; secondly, such films can cause solution-transfer creep[8,17-20]; thirdly, water can diffuse into the sodium chloride lattice[21,22], possibly affecting dislocation dynamics. To investigate these possibilities, we conducted two series of experiments.

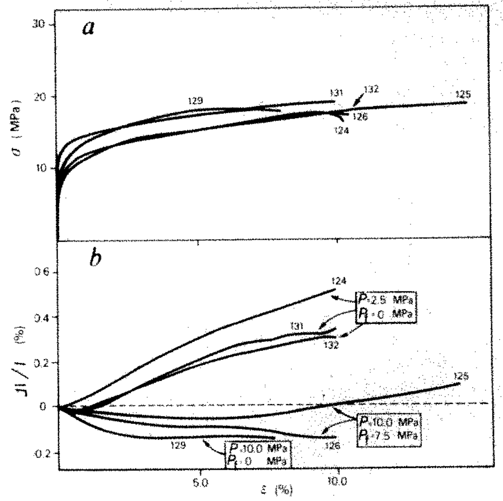


Fig. 1 *a*, Differential stress (σ) and *b*, volume change ($\Delta V/V$) plotted against strain (ε) for the pre-relaxation stage of the deformation experiments on natural (Asse) samples. $T = 150\,°\mathrm{C}$; $\dot{\varepsilon} = 3\times10^{-5}\,\mathrm{s}^{-1}$. The confining pressure (P) and pore fluid (brine) pressure (P_f) used in each test are indicated. Note that samples deformed at $P = 2.5$ MPa show substantial dilatancy (volume increase) compared with those deformed at $P = 10$ MPa.

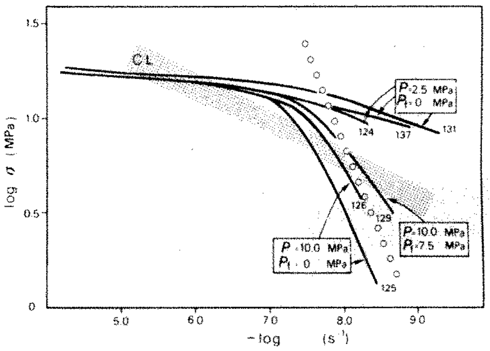


Fig. 2 Differential stress (σ) plotted against strain rate ($\dot{\varepsilon}$) (solid curves) for the stress relaxation stage of the experiments shown in Fig. 1. Note the dramatic weakening effect seen in samples deformed in non-dilatant conditions ($P = 10.0$ MPa), at strain rates of $<10^{-7}\,\mathrm{s}^{-1}$. This low-strain-rate behaviour can be described by a power-law creep equation, $\dot{\varepsilon} = A\sigma^n$, where the stress exponent (or stress sensitivity of strain rate, n) has a value of 1-2. Samples deformed in dilatant conditions ($P = 2.5$ MPa) are stronger than predicted by previously existing creep laws[1,3,4,25] (stippled area, Cl) but exhibit similar stress sensitivity (n values) at the lower strain rates. Note that the weakening behaviour observed in non-dilatant conditions ($P = 10.0$ MPa) agrees well with predictions based on the solution-precipitation creep model given in the text (equation (1)). This model was applied to the recrystallized samples using $d = 3$ mm, $a = 44$ (ref. 18) and $C = 8\times10^{-8}$ m (ref. 8). The diffusivity parameter B was estimated from published data[8].

In the first series, we used natural rock salt samples from the Asse Mine (West Germany). The material consists of a very pure (99%) halite rock, with a grain size of ~5 mm (Fig. 3*a*) and a dynamically recrystallized microsctructure[8,9]. The water content is ~0.05 wt%[23], up to half of which is present in fluid inclusions at grain boundaries (Fig. 3*b*, *i*). Right-cylindrical samples, 70 mm in diameter and 150 mm long, were machined from this material, and deformed at 150 °C at a constant strain rate of $\sim10^{-5}\,\mathrm{s}^{-1}$, followed by stress relaxation[8,24]. A dilatometric triaxial testing machine was used[8]. The pre-relaxation stress-strain behaviour was similar for all samples (Fig. 1), but clear differences were observed in dilatancy and stress relaxation (Figs 1, 2). Two types of behaviour can be distinguished: (1) Samples deformed at a confining pressure of

---

[*] Lamont-Doherty Geological Observatory, Palisaes, New York 10964, USA.

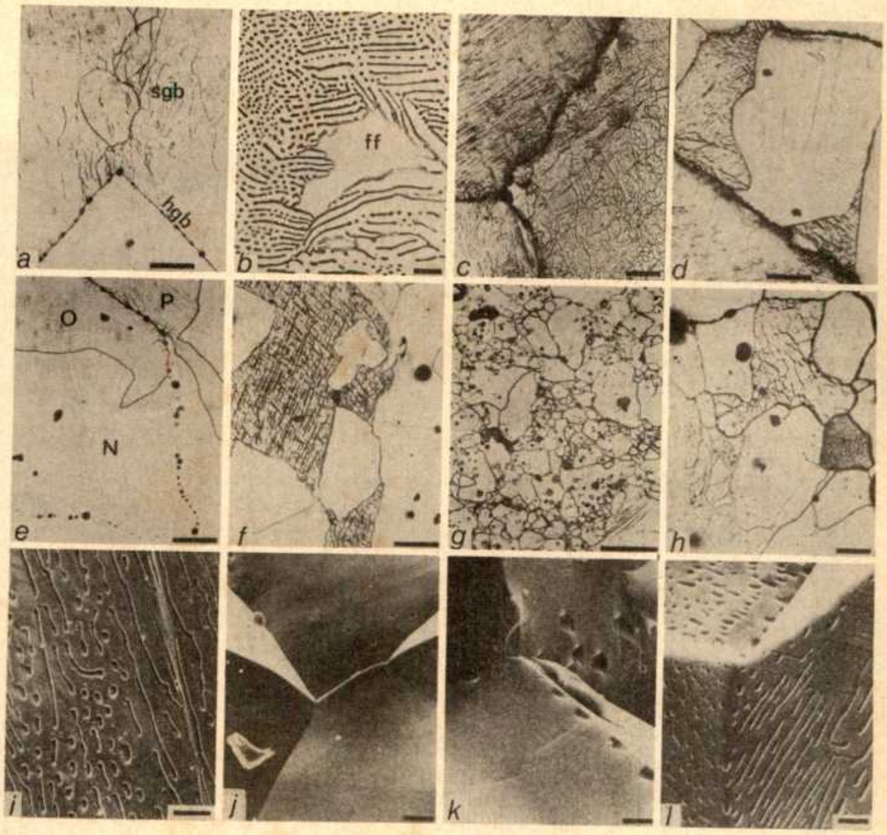


Fig. 3 *a*, Reflected-light micrograph of a polished and etched section of the natural starting material (from the Asse Mine) used in the present experiments. A grain containing subgrain boundaries (sgb) with a few degrees misorientation across them can be seen to be replaced by a newer grain, with fluid inclusions developed in the high-angle grain boundary (hgb) separating the two grains. Scale bar, 0.1 mm. *b*, Transmitted-light micrograph of the Asse salt shown in *a*, showing the typical structure of a high-angle boundary in plan view (see *i*). Note the array of fluid inclusions and fluid film segments (ff). These generally contain gas and/or brine. Scale bar, 0.1 mm. *c*, Reflection micrograph showing typical microsctructure of Asse salt samples deformed in dilatant conditions ($P = 2.5$ MPa). Note the densely developed slip-band structures, crude subgrains and dilated grain boundaries. Scale bar, 0.1 mm. *d*, Reflection micrograph illustrating the typical recrystallized microstructure of Asse samples deformed in non-dilatant conditions ($P = 10.0$ MPa). The micrograph shows a deformed grain with heavily developed dislocation substructure being replaced by a new grain with no visible substructure. The boundaries of new grains such as these can often be shown to contain thin brine films. Scale bar, 0.1 mm. *e*, Lobate new grain (N) consuming heavily deformed old grains (O, P) in a recrystallized Asse sample deformed in non-dilatant conditions. Note that gas-bearing inclusions in the boundary between grains O and P have been overgrown by the new grain N to form a 'ghost' grain boundary. Scale bar, 0.1 mm. *f*, Reflection micrograph showing dynamic recrystallization in a rock salt sample from Avery Island deformed to ~40% strain at $T = 70$ °C and $P = 20.4$ MPa, with differential stresses in the range 10.3–20.4 MPa. Again, note how new grains with relatively little substructure are growing at the expense of the old, heavily substructured grains. Scale bar, 0.1 mm (sample courtesy of Dr F. D. Hansen). *g*, Reflected-light micrograph of deformed sample of artificially prepared salt (grain size ≈80 μm; see Fig. 4 for mechanical data). Note the fine, equiaxial, largely recrystallized grain structure. The dark spots are pores emerging at the surface of the polished section. Scale bar, 0.1 mm. Neutron diffraction methods have shown that this sample has no significant crystallographic preferred orientation (H. G. Brokmeier, personal communication). *h*, Fine-grained, recrystallized microstructure in a sample from the salt glacier described in ref. 14. Compare this with *a*, *d*, *f* and *g*. Scale bar, 0.1 mm. *i*, Scanning electron micrograph of a mechanically parted grain boundary in a sample of the Asse salt used as starting material in our experiments on natural material (see *b*). Note the tendency of the grain-boundary fluid inclusions to develop negative crystal forms. Scale bar, 0.04 mm. *j*, Scanning electron micrograph of grain surface in a recrystallized Asse sample (see *d*), one month after the end of deformation in non-dilatant conditions. Note the smoothness of these surfaces. Such grain boundaries can be shown to contain brine by disrupting them *in situ* under the transmission optical microscope. The brine film thickness has been shown to be <100 nm by measuring triple-junction void diameters from scanning electron micrographs. Scale bar, 0.1 mm. *k*, *l*, Scanning electron micrographs of grain surfaces in a recrystallized Asse sample, tested in identical conditions to the sample shown in *j*, but observed 1 yr after the test ended. Note the clear changes in morphology. It is inferred that the observed arrays of inclusions formed by the breakdown of originally subcontinuous brine films, presumably by means of a solution–precipitation process driven by surface energy. This also offers an explanation for the grain-boundary inclusion arrays seen in the Asse starting material (*b*, *i*). Scale bars, 0.02 mm.

2.5 MPa (and without added brine) show marked dilatancy during the first stage of deformation. During relaxation at strain rates of $>10^{-7}\,\mathrm{s}^{-1}$, the differential stress is relatively insensitive to strain rate (Fig. 2). Microstructural analysis of the samples reveals densely developed deformation-band structures and crude subgrains (Fig. 3*c*). Grain boundaries are commonly dilated and fluid inclusions are disrupted (Fig. 3*c*). The overall behaviour agrees well with previous findings: work hardening is accompanied by some climb-controlled recovery[1,25].

(2) Samples deformed at 10 MPa confining pressure (with and without brine added) show little or no dilatancy during the first stage of the test. Furthermore, during stress relaxation there is a marked weakening below strain rates of $10^{-7}\,\mathrm{s}^{-1}$ (Fig. 2). In these samples, the grain size is reduced to about half that of the starting material. Optical microscopy shows deformed grains being replaced by new, recrystallized grains (Fig. 3*d*, *e*, *f*). The samples are ~70% recrystallized. New grain boundaries contain subcontinuous brine films (Fig. 3*i*–*l*) which are <100 nm thick.

The most likely explanation of the type (2) behaviour is as follows. During the first stage of the experiments, brine present in grain-boundary voids (Fig. 3*b*) is transformed into subcontinous brine films[26] which enhance grain boundary mobility. The grain boundaries then migrate, and the sample recrystallizes into a finer aggregate with lower dislocation density, and a network of brine-filled grain boundaries. The low values of the stress exponent n at strain rates below $10^{-7}\,\mathrm{s}^{-1}$ (Fig. 2) suggest that the dominant deformation mechanism changes from dislocation to solution-transfer creep[8,18,19]. This is supported by the finding (Fig. 2) that the observed weakening behaviour agrees well with the following theoretical model for diffusion-controlled solution-transfer creep[8]:

$$\dot{\varepsilon} = \frac{ABC\sigma}{kTd^3} \qquad (1)$$

where $\dot{\varepsilon}$ is the strain rate, σ is the applied stress, A is a constant, B is a temperature-dependent diffusivity term, C is a grain-boundary structure parameter, k is the Boltzmann constant, T is the absolute temperature and d is the grain size.

At low strain rates, the samples showing type (2) behaviour are much weaker than expected from previous data (see Fig. 2), the weakening effect becoming more pronounced towards lower strain rates. The absence of this weakening in type (1) samples

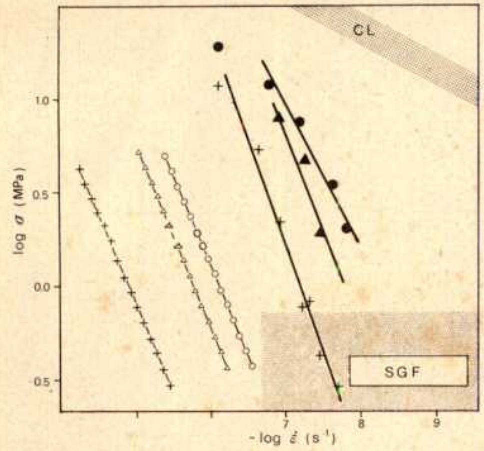


Fig. 4 Log σ plotted against log $\dot{\varepsilon}$ for three experiments on wet synthetic salt (solid lines), compared with the behaviour expected on the basis of previously existing creep laws[1,3,4,25] (stippled area, CL). The broken lines represent the behaviour predicted, for each of the three experiments reported, using the solution–precipitation creep model (eqn 1). Experiments: ●, $T = 20\,°C$; $d = 0.2$ mm; $P = 21$ MPa; $P_f = 3$ MPa; ▲, $T = 70\,°C$; $d = 0.2$ mm; $P = 21$ MPa; $P_f = 3$ MPa; +, $T = 20\,°C$; $d = 0.08$ mm; $P = 5.1$ MPa; $p_f = 0.1$ MPa. Model: ○, $T = 20\,°C$; $d = 0.2$ mm; △, $T = 70\,°C$; +, $T = 20\,°C$; $d = 0.08$ mm. Note the excellent agreement regarding stress sensitivity of strain rate ($n \approx 1$). The model was applied using $A = 44$ (ref. 18) and $c = 8 \times 10^{-8}$ m (ref. 8). The value of B was estimated from published data[8]. Finally, note that the experimental results reported in this diagram are almost directly relevant to the σ–$\dot{\varepsilon}$ conditions (stippled area, SGF) associated with the flow of wet salt glaciers in nature[13] ($d = 0.2$ mm).

(those samples deformed dry at pressure allowing dilatancy) can be explained by grain-boundary microcracking. This almost certainly leads to evaporation of the grain-boundary fluid.

In the second series of experiments we used artificial, fine-grained samples expected to exhibit accelerated solution-transfer creep effects (see equation (1)). These were prepared by dry compaction and annealing of fine-grained analytical-grade NaCl. The experiments were carried out in non-dilatant conditions using samples of 30 mm diameter and 60 mm length containing ~1% added brine[8]. A strain-rate stepping technique was used. The results indeed show the expected 'low n' behaviour (Fig. 4) and grain size dependence. Microstructural investigations show an equiaxial, recrystallized fabric, with little or no dislocation substructure (Fig. 3g), and no preferred lattice orientation. Grain boundaries were fluid-filled, with thicknesses of the same order as the recrystallized Asse samples. Most of the fluid was located in voids. Taken together, these observations provide evidence[20] that diffusive mass transfer was the main deformation mechanism, accompanied by grain-boundary sliding and migration. Comparison of the experimental data with the model of equation (1), applied using reasonable values of the parameters involved, gives good agreement for the stress exponent n (Fig. 4). Note also that the present experiments are almost directly relevant to the conditions found in salt glaciers (Figs 3h, 4), and show the weak behaviour reported in ref. 13.

Natural rock salt invariably contains small amounts of brine[16], part of which resides in grain-boundary fluid inclusions and films. Both sets of experiments reported here show that grain-boundary brine can have large effects on rheology and deformation mechanisms, by promoting dynamic recrystallization and solution transfer. In our opinion the weakening reported here has not been observed by previous workers on natural salt, either because the lowest strain rates reached were too high (Fig. 2), or because the water inherent in the samples was lost before testing. Note that most previous work has been performed in dilatant conditions. It is clear that in future long-term creep work, careful consideration should be given to recrystallization and grain-boundary fluid effects.

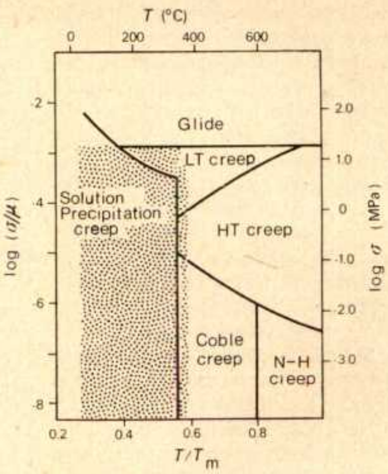


Fig. 5 Deformation mechanism map for halite (after ref. 28), modified to include solution-precipitation creep as described by the model given in the text (equation (1)). Values of the parameters A, B and C were chosen as for Figs 2 and 4. The map was constructed for a grain size of 1 mm and intergranular fluid pressure of 10 MPa. The extent of the solution-precipitation or solution-transfer creep field towards higher temperatures is assumed to be limited by the liquid–vapour phase transition. Stippling indicates the range of conditions relevant for halokinesis and radioactive waste repositories (the latter ranging up to only 200 °C). LT, low-temperature; HT, high-temperature; N–H, Nabarro–Herring.

Keeping in mind that our natural (Asse) starting material is relatively dry, weakening by water should occur in a wide range of salt rocks during natural deformation[17,27]. A rough extrapolation of our data to the grain sizes, tempertures, and differential stresses[7] expected during natural salt deformation suggests that brine-related effects should become important at strain rates of $\sim 10^{-10}\,s^{-1}$ (just below those accessible in the laboratory), and we propose that most natural salt deformation occurs in the transition zone between the dislocation-creep and solution-transfer fields. We further conclude that the extrapolation of previous flow laws to predict long-term behaviour is questionable, unless the effects reported here are also taken into account. Construction of a deformation mechanism map for salt (Fig. 5), incorporating solution–precipitation creep (equation (1)), shows that brine effects can be expected throughout a very wide range of natural conditions, including those relevant to radioactive waste repositories and storage caverns.

We thank H. Gies, C. J. Talbot and F. D. Hansen for supplying samples, C. J. Peach and J. L. Liezenberg for discussions and assistance with experimental work, J. N. Boland for help with electron microscopy, H.-G. Brokmeier for measuring preferred lattice orientations, and the workshop staff of the HPT laboratory at Utrecht for technical assistance. The work was performed under contract with the Netherlands Ministry of Economic Affairs and the Commission of the European Communities (contract WAS-153-80-7-N) in connection with their research programmes on radioactive waste management and storage.

Received 25 June; accepted 30 September 1986.

1. Carter, N. L. & Hansen, F. D. *Tectonophysics* **92,** 275–333 (1983).
2. Jackson, M. & Talbot, C. J. *Bull. geol. Soc. Am.* **97,** 305–323 (1986).
3. Langer, M. In *Proc. 1st Conf. on the Mechanical Behaviour of Salt* (eds Hardy, H. R. & Langer, M.) 201–240 (Trans-Tech, Clausthel-Zellerfeld, 1984).
4. Herrmann, W., Wawersik, W. R. & Lauson, H. S. *Sandia natn. Lab. Rep.* No. SAND80-2712 (1980).
5. Wawersik., W. R., Herman, W., Montgomery, S. T. & Lauson, S. in *Proc. int. Symp. Rock Mechanics related to Caverns and Pressure Shafts* (ed. Wittke, W.) 1345–1356 (Balkema, Amsterdam, 1982).
6. Skrotzki, W. & Welch, P. *Tectonophysics* **99,** 47–61 (1983).
7. Carter, N. L., Hansen, F. D. & Senseny, P. E. *J. geophys. Res.* **87,** 9289–9300 (1982).
8. Spiers, C. J., Urai, J. L., Lister, G. S., Boland, J. N. & Zwart, H. J. in *Nuclear Science and Technology* (Office for Official Publs. of the European Communities, Luxembourg).

9. Urai, J. L., Spiers, C. J., Liezenberg, J. L., Franssen, R. C. & Peach, C. J. *Geol. Minb.* (in the press).
10. Müller, H. G. *Z. Phys.* **96,** 279-327 (1935).
11. Friedman, M., Dula, W. F., Gangi, A. F. & Gazonas, G. A. in *Proc. 1st Conf. on the Mechanical Behaviour of Salt* (eds Hardy, H. R. & Langer, M.) 19-36 (Trans-Tech, Clausthel-Zellerfeld, 1984).
12. Guillope, M. & Poirier, J. P. *J. geophys. Res.* **84,** 5557-5567 (1979).
13. Talbot, C. J. & Rogers, E. A. *Science* 208-395 (1980).
14. Talbot, C. J. *Spec. Pap. geol. Soc. Am.* **9,** 173-184 (1981).
15. Wenkert, D. D. *Geophys. Res. Lett.* **6,** 523-526 (1979).
16. Roedder, E. A,. *Miner.* **69,** 413-439 (1984).
17. Urai, J. L. *Tectonophysics* **96,** 125-157 (1983).
18. Raj, R. *J. geophys. Res.* **87,** 4731-4739 (1982).
19. Lehner, F. K. & Bataille, J. *Pure appl. Geophys.* **122,** 53-85 (1984).
20. Raj, R. & Ashby, M. F. *Metall. Trans.* **2,** 1113-1127 (1971).
21. Ruhenbeck, C. *Z. Physik.* **207,** 446-469 (1967).
22. Green, A. K. & Bauser, E. *J. appl. Phys.* **39,** 2769-2779 (1968).
23. Jockwer, N. thesis, Tech. Univ. Clausthal (1981).
24. Spiers, C. J., Urai, J. L. & Lister, G. S. in *Proc. 2nd Conf. Mechanical Behaviour of Salt,* Hannover (in the press).
25. Heard, H. C. *Am. geophys. Un. Geophys. Monogr.* **16,** 191-209 (1972).
26. Urai, J. L., Means, W. D. & Lister, G. S. *Am. geophys. Un. Geophys. Monogr.* **36,** 161-199 (1986).
27. Urai, J. L. *Tectonophysics* **120,** 285-317 (1985).
28. Albrecht, H. & Hunsche, U. *Fortschr. Miner.* **58,** 212-247 (1980).

Diagenesis of US Gulf Coast shales

K. Pye*, D. H. Krinsley† & J. H. Burton†

* Department of Earth Sciences, University of Cambridge,
Downing Street, Cambridge CB2 3EQ, UK
† Department of Geology, Arizona State University, Tempe,
Arizona 85287, USA

The burial diagenesis of mudstones is of major interest in petroleum geology[1-3]. Key questions which remain incompletely answered include the relative importance of temperature, time, changes in pore-water chemistry and organic maturation as driving mechanisms, and the relationship between diagenesis in mudstones and neighbouring sandstones[4-7]. We reconsider here the nature and timing of diagenesis in Gulf Coast Tertiary mudstones using data obtained by backscattered electron microscopy (BSEM) and energy-dispersive X-ray analysis (EDS) of cuttings from two Texas wells. Our observations confirm the mineralogical trends with depth reported by Hower *et al.*[1] and suggest new evidence concerning the timing and causes of the changes. The most important mineralogical changes occurred in the zone of organic matter decarboxylation (zone IV[8]). We show that the sequence of burial diagenetic events can be established quickly and reliably by BSEM examination of shale cuttings, and demonstrate the use of foram tests and their authigenic mineral infillings as indicators of diagenesis in soft, fine-grained shales.

The Tertiary mudstones and interbedded sandstones of the USA Gulf Coast represent one of the world's most intensively studied sedimentary sequences. Many fundamental concepts concerning the nature, timing and consequences of shale diagenesis were developed through work in this area[1,5,9-12]. Of particular interest has been the transformation of smectite in mixed-layer illite-smectite clays, involving an apparent increase in the proportion and ordering of illite layers with burial. It has been suggested that release of water during this transformation may be important in flushing out hydrocarbons from source rocks, in transporting ions for the formation and/or dissolution of cements in neighbouring reservoir sandstones, and in leading to over-pressuring and hydro-fracturing[3,9,13].

Several important questions have not been resolved in the classic work by Hower and his associates[1,11,14,15]. In particular, the sequence of mineralogical changes and the driving mechanism behind the diagenetic transformations were not positively identified. Decomposition of detrital potassium feldspar and mica during burial was suggested to provide the additional K^+ and Al^{3+} required for the transformation of smectite to illite[1], with release of Si^{4+}, Fe^{2+} and Mg^{2+} (reprecipitated as chlorite and quartz). It was suggested that substitution of Al^{3+} for Si^{4+} in the tetrahedral layers of the smectite was the principal factor allowing fixation of K^+ in the interlayer positions by increasing the layer negative charge. The reasons for the decomposition of K-feldspar and mica were not identified. Hower *et al.* also observed a gradual loss of calcite over the depth interval in which the clay transformations occurred (2,500-3,700 m), but no causal connection was suggested.

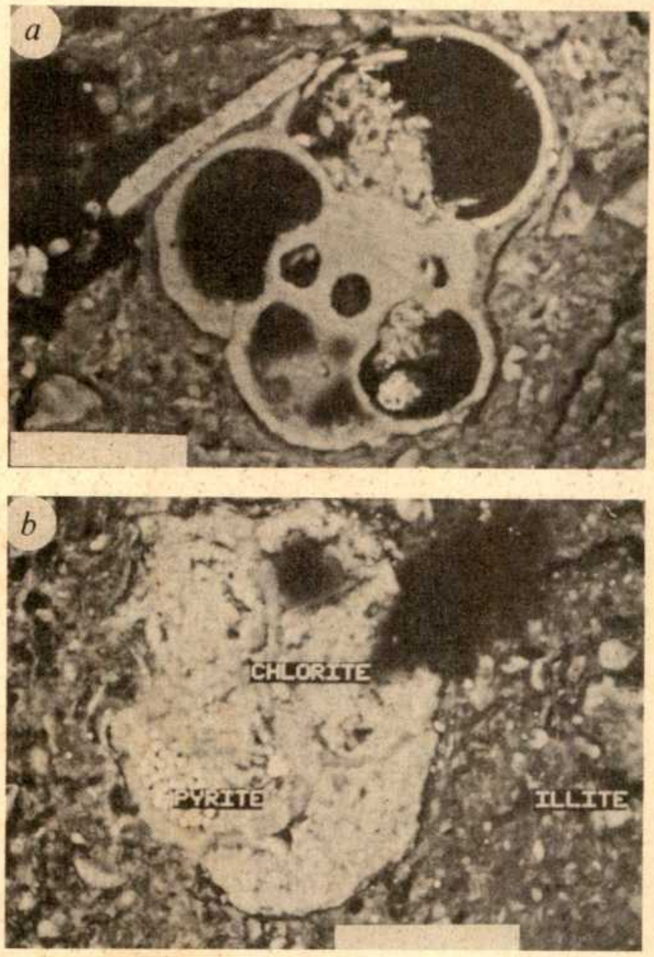


Fig. 1 BSEM micrographs of polished shale cuttings from C.W.R.U. Gulf Coast Well No. 6 showing: *a*, a calcite foram test partly infilled by authigenic chlorite (depth 2,680 m); and *b*, a foram completely infilled and replaced by chlorite (depth 3,415 m). The surrounding matrix consists mainly of illite-smectite and fine-grained quartz (scale bar, 50 μm).

De Segonzac[16] postulated that pressure is a major factor affecting clay transformations, but Perry and Hower[11] and Heling[17] concluded that temperature is more important. Other authors have stressed the role of pore-water chemistry, particularly K^+ availability[12,18]. Iron reduction[7,19] and organic matter maturation, leading to marked changes in shale porewater pH[5,6,20], have also been suggested to be important controls.

Previous Gulf Coast shale studies have relied mainly on 'bulk' methods of mineralogical and chemical analysis, particularly X-ray diffraction (XRD). No direct observations of textural relationships between different minerals with depth have been made, and diagenetic processes such as dissolution, precipitation and recrystalization have mainly been inferred from changes in relative mineral, elemental and isotopic abundance in different grain size fractions. We considered that useful information might be provided by examining polished shale cuttings using BSEM and EDS[21]. Cuttings from two south Texas wells were examined. The first sample suite spanned the depth interval 1,200-3,700 m in Case Western Reserve University Gulf Coast Well No. 6, previously studied by Hower *et al.*[1]. The second

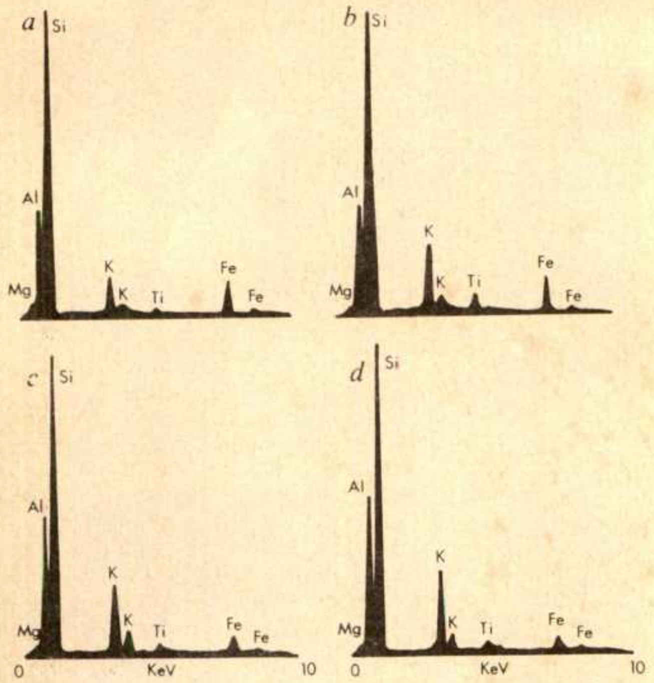


Fig. 2 Typical EDS spectra of illite–smectite aggregates showing observed changes in elemental peak heights with depth in Mobil Halls Bayou Wildcat No. 1. Samples are from depths of: *a*, 2,232 m; *b*, 3,650 m; *C*, 4,485 m; *d*, 5,530 m.

suite came from Mobil Halls Bayou Ranch No. 1 Wildcat, Galveston County. Both wells penetrate virtually the entire thickness of the Frio and Anahuac Formations which were deposited as a rapidly subsiding sediment wedge on the north side of the Gulf Coast Tertiary Basin[22,23]. The upper Frio consists of interbedded terrestrial and marginal marine sandstone and shale units underlain by >2,000 m of undercompacted slope and basinal mudstones. The Anahuac represents a transgressive, predominantly marine shale wedge which onlaps the Frio.

Shale cuttings from selected depth intervals were oriented and impregnated in epoxy for scanning electron microscope (SEM) examination transverse to the bedding. Due to their fine-grained, soft nature, many specimens were polished by hand to minimize surface damage. After carbon coating the specimens were examined in an ISI SEM equipped with a solid-state BSE detector and Kevex EDS analyser. Cuttings from each depth interval were also analysed by XRD.

Useful textural and chemical data were obtained from all but the finest-grained cuttings. Even in specimens composed largely of clay, clear indications of diagenetic reactions are provided by authigenic mineral aggregates within or replacing foram tests. In Gulf Coast Well No. 6, calcite foram tests are abundant above 2,750 m (Fig. 1*a*). The tests are partly filled by authigenic pyrite, kaolinite and chlorite, with some unoccupied pore space. Between 2,750 and 3,400 m the remaining voids are filled by authigenic chlorite and/or illite (Fig. 1*b*), while the tests themselves are dissolved. In some cases, the chambered structure of the tests is preserved by chlorite replacement, indicating that infilling of the intra-test voids occurred earlier than dissolution of the test walls. In the same depth interval, zoned authigenic carbonate rhombs, which were formed in the sulphate reduction and fermentation zones (zones II and III[8]), are also dissolved and replaced by iron-rich chlorite. These observations indicate a significant reduction in shale porewater *p*H in the temperature range 60–90 °C (see ref. 1 for down-hole temperature data).

Although the Frio mudstones contain many clay particles which cannot be resolved individually using BSEM, much of the illite–smectite clay (identified by XRD) occurs as aggregates up to 20 μm in size which may be compacted floccules. The aggregates are sufficiently mono-mineralic to allow semi-quantitative X-ray analysis of their chemical composition. The EDS spectra in Fig. 2 illustrate the general nature of the changes in composition of the illite–smectite aggregates with depth in the Mobil well. Similar trends were observed in C.W.R.U. Well No. 6. The illite–smectite above 2,500 m has a high Si/Al peak height ratio and a low K/Fe ratio, reflecting an abundance of smectite components. With increasing depth there is an increase in Al and K and a relative reduction in Si and Fe. This is in agreement with previous wet chemical data obtained from the fine-clay fractions[1] and corresponds with an increase in the proportion of illite components indicated by XRD[1]. In the Mobil well, the first change in composition of the illite–smectite indicated by EDS occurs at 3,620 m. At this depth, K abundance increases whereas Fe and Al abundance remains unchanged. With increasing depth there is a marked increase in Al and decrease in Fe while the K content remains almost constant. These observations suggest that the conversion of smectite to illite (if some form of solid-state transformation is occurring) is not dependent on tetrahedral substitution of Al^{3+} for Si^{4+}. However, it is possible that the creation of octahedral layer charge, due to the reduction of Fe^{3+}, during passage of the sediment through diagenetic zones II, III and IV, contributes to the fixation of interlayer K^{+}, as previously suggested by Eslinger *et al.*[24].

The infilling and replacement of some foram tests by illite indicates that this mineral is at least partly neoformed by direct precipitation from pore fluids.

K-feldspar was not present in large amounts in any of the cuttings studied, but bulk rock XRD analyses and SEM evidence indicated some reduction in abundance at slightly greater depth than carbonate dissolution, suggesting that carbonate may have acted as a temporary buffer which retarded dissolution of feldspar.

It has been suggested that thermal decarboxylation of kerogen in mudstones is the most probable mechanism whereby acid pore waters are generated at depth[20,25]. This may occur where production of CO$_2$ by decarboxylation exceeds iron reduction[7]. In the Gulf Coast, excess decarboxylation is favoured by a relatively high geothermal gradient and a rapid rate of burial. However, the Frio shales contain on average only 0.28% organic matter, and could not generate sufficient CO$_2$ to account for all of the secondary porosity observed in Frio sandstones plus the loss of carbonate originally present in the shales[5]. Aliphatic and other organic acid anions are also produced during decarboxylation[20,26], but a similar mass balance problem arises[5]. A further possibility is that, once initiated, the smectite–illite transformation generates residual acid, proton-donating water which may contribute to a lowering of *p*H[27,28].

From the textural and chemical data provided by BSE and EDS we conclude that the sequence of diagenetic changes was (1) reduction of *p*H due to excess decarboxylation during simultaneous iron reduction; (2) dissolution of carbonate; (3) dissolution of detrital feldspars and micas; (4) build-up of negative octahedral layer charge due to reduction of Fe^{3+} in mixed-layer clays; (4) fixation of interlayer K^{+} and substitution of Al^{3+} for tetrahedral Si^{4+}, the additional K and Al being derived from dissolved detrital silicates; (5) precipitation of neoformed illite and chlorite in pores, and partial replacement of kaolinite by chlorite. Some chlorite and/or berthierine[29] was precipitated within foram tests before the onset of carbonate dissolution in zone IV.

The pattern of shale diagenesis found in the Gulf Coast differs from that found in some other sedimentary basins, and it may be unwise to regard it as a general model. The Gulf Coast sediments initially contained a large iron-bearing smectite and mixed-layer clay component, were buried relatively rapidly, and experienced high thermal gradients. Consequently they bear an important imprint from diagenetic zone IV. In contrast, many British Jurassic mudstones initially contained little smectite or mixed-layer illite-smectite compared with kaolinite and degraded micas, and experienced low burial rates, with the result that the rocks bear a dominant mineralogical imprint from diagenetic zones II and III[30-32]

This research was supported by NATO grant RG 84/0039, the Petroleum Research Fund, the American Chemical Society, and the American Association of Petroleum Geologists

Received 9 June, accepted 10 October 1986

1 Hower, J, Eslinger, E V, Hower, M E & Perry, E A *Bull geol Soc Am* **87**, 725-737 (1976)
2 Foscolos, A E & Powell, T G in *International Clay Conference 1978* (eds Mortland, M M & Farmer, V C) 262-270 (Elsevier, Amsterdam, 1978)
3 Bruce, C H *Am Ass petrol Geol Bull* **68**, 673-683 (1984)
4 Boles, J R & Franks, S G *J sedim Petrol* **49**, 55-70 (1979)
5 Lundgard, P D, Land, L S & Galloway, W E *Geology* **12**, 399-402 (1984)
6 Surdam, R C, Boese, S W & Crossey, L J *Am Ass petrol Geol Mem* **37**, 127-149 (1984)
7 Curtis, C D *Phil Trans R Soc* A **315**, 91-105 (1985)
8 Curtis, C D *Geol Soc Lond Spec Publ* 113-125 (1985)
9 Powers, M C *Am Ass petrol Geol Bull* **51**, 1240-1254 (1967)
10 Burst, J F *Am Ass petrol Geol Bull* **53**, 73-93 (1969)
11 Perry, E A & Hower, J *Am Ass petrol Geol Bull* **56**, 2013-1021 (1972)
12 Morton, J P *Bull geol Soc Am* **96**, 114-122 (1985)
13 Foster, W R & Custard, H C *Am Ass petrol Geol Bull* **64**, 708 (1980)
14 Aronson, J L & Hower, J *Bull geol Soc Am* **87**, 738-744 (1976)
15 Yeh, H-W & Savin, S M *Bull geol Soc Am* **88**, 1321-1330 (1977)
16 De Segonzac, D *Sedimentology* **15**, 281-346 (1970)
17 Heling, D *Sedimentology* **21**, 463-472 (1974)
18 Roberson, H E & Lahann, R W *Clays Clay Miner* **29**, 129-135 (1981)
19 Coleman, M L *Phil Trans R Soc* A **315**, 39-56 (1985)
20 Curtis, C D *Am Ass petrol Geol Bull* **67**, 380-383 (1983)
21 Pye, K & Krinsley, D H *J sedim Petrol* **54**, 877-888 (1984)
22 Loucks, R G, Bebout, D G & Galloway, W E *Trans Gulf Coast Ass geol Soc* **27**, 109-120 (1977)
23 Galloway, W E, Hobday, D K & Magara, K *Bur Econ Geol Univ Texas, Austin, Rep Investign* **122**, 1-78 (1982)
24 Eslinger, E Highsmith, P, Albers, D & DeMayo, B *Clays Clay Miner* **27**, 327-338 (1979)
25 Schmidt, V & McDonald, D A *Soc Econ Paleontol Miner Spec Publ* **26**, 175-207 (1979)
26 Carothers, W W & Kharaka, Y K *Am Ass petrol Geol Bull* **62**, 2441-1453 (1978)
27 Mortland, M M *Trans 9th int Cong Soil Sci* **1**, 691-699 (1968)
28 Johns, W D & Shimoyama, A *Am Ass petrol Geol Bull* **56**, 2160-2167 (1972)
29 Ahn, J H and Peacor, D R *Clays Clay Miner* **33**, 228-236 (1985)
30 Irwin, H, Curtis, C D & Coleman, M L *Nature* **269**, 209-213 (1977)
31 Morris, K A *J geol Soc Lond* **137**, 157-175 (1980)
32 Pye, K & Krinsley D H *Geol Mag* **123**, 191-203 (1986)

Do continental shelves export organic matter?

G. T. Rowe*, S. Smith*, P. Falkowski*, T. Whitledge*, R. Theroux†, W. Phoel‡ & H. Ducklow§

* Oceanographic Sciences Division, Department of Applied Science, Brookhaven National Laboratory, Upton, New York 11973, USA
† Northeast Fisheries Center, NMFS/NOAA Woods Hole, Massachusetts 02543, USA
‡ Northeast Fisheries Center, NMFS/NOAA Sandy Hook, New Jersey 07732, USA
§ University of Maryland, Center for Environmental Studies, Cambridge, Maryland 21613, USA

It has been suggested that biological production and consumption of organic matter is not balanced in coastal marine ecosystems[1,2]. If, as suggested, 90% of the phytoplankton produced during the spring bloom period were exported, excess organic carbon would be sequestered on the continental slope below the permanent thermocline. Here we summarize the shelf-edge exchange processes (SEEP) experiment, designed to test the export hypothesis. The absence of a positive imbalance in the organic carbon budget, reinforced by modest sediment deposition and biomass on the continental slope, led us to reject the concept. Only a small fraction of continental shelf phytodetritus is exported; that not consumed in the spring is for the most part used on the continental shelf during the ensuing stratified season The original hypothesis failed to recognize the contribution of pelagic microbial consumption and the lag in coupling between seasonal production and consumption processes.

Table 1 Moored sediment trap (MST) flux

	Date	Mass flux (mg m^{-2} d^{-1})	C (mg^{-2} d^{-1})	N (mg m^{-2} d^{-1})
MST 1	25 Feb - 5 March 1984			
70 m o b *		1,240 ± 198	54 ± 6	6 ± 1
10 m o b		5,450 ± 467	199 ± 44	20 ± 7
MST 2	5-31 March 1984			
70 m o b		1,810 ± 239	110 ± 28	12 ± 5
10 m o b		3,815 ± 418	159 ± 18	19 ± 4
MST 3	4-7 April 1984			
70 m o b		2,298 ± 240	140 ± 12	6 ± 4
10 m o b		2,748 ± 624	107 ± 16	1 ± 1

Data from SEEP I (150 m, west line locality W12 shown on Fig 1) Values are mean ± s d
* m o b, metres off bottom

The data we present come from a restricted area of the continental shelf of the eastern United States south of New England, which extends from longitudes 70°30′ to 72°30′ W, and includes the area between the coastline and the outer edge of the continental shelf from Martha's Vineyard south to Long Island (Fig 1) Two transects of moorings were instrumented with current meters, transmissometers and internally recording fluorometers[3,4] between February and April 1984 Sediment traps[5] were moored at locality W12 (Fig 1) during this period, and at the deeper moorings on the east line[6] throughout the remainder of the year Estimates of standing stocks and a variety of biological rate measurements were made during 24-48 h periods in shelf and slope water on both lines (Fig 1) between late February and mid-April 1984 Measurements over similar periods were made during July 1983 On the basis of the original hypothesis these data should have shown that annual net primary production appreciably exceeded consumption on the continental shelf, that biogenic particles or plant pigments had a net offshore flux, and that intensified organic carbon deposition was reflected in the sediments or the biota on the continental slope

The main aim of our work was to study the shelf carbon budget and to search for an imbalance betweeen primary production and heterotrophic use within the study area (Fig 1) Primary productivity was estimated by measuring ^{14}C-labelled assimilation during incubations aboard ship[7], and the consumption of phytoplankton carbon was estimated by summing respiration and secondary production of macrozooplankton[8-10], the benthic community[11,12] and pelagic bacteria[13,14] Microzooplankton were not included Zooplankton grazing was estimated from gut clearance rates[10] and known relationships between feeding rates, body size and food concentration[9] Bacterial respiration and production were based on labelled thymidine incorporation[13] and NH$_4$ regeneration rate measurements using ^{15}N-labelled substrates[14] Use of organic matter by the benthic biota was assumed to be by respiration and secondary production of the macrofauna[15] Respiration, the sum of anaerobic microbial metabolism and oxygen use, was estimated from oxygen use *in situ* in benthic chambers placed on the bottom by scuba divers and by ship-board incubations of the sediment cores[12] Sulphate reduction, nitrate reduction and organic carbon accumulation in shelf sediments may also contribute to carbon fluxes[12], but were not included here Although total benthic standing stocks were measured[12], we include here only the secondary production of the macrofauna based on established production to biomass ratios[15,16] and measured biomass of the macrofauna[17,18] Secondary production of the meiofauna and benthic bacteria were not included

The critical issue is the comparison of phytoplankton production rate with the total rate of consumption of organic matter

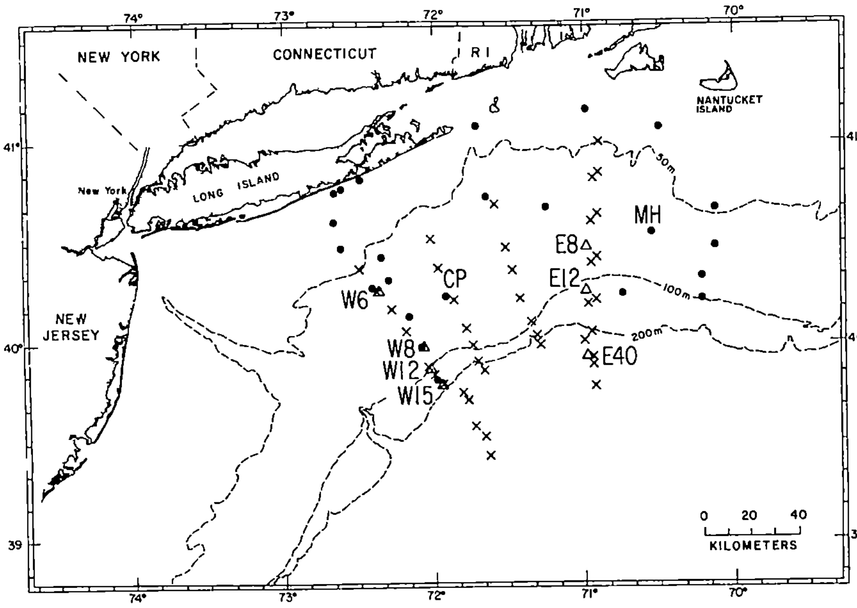


Fig 1 Location of SEEP experiment, east and west lines Time series were conducted at the triangles Sediment traps were located at W12 ●, Locations of sediment oxygen demand measurements, ×, locations of core or grab samples

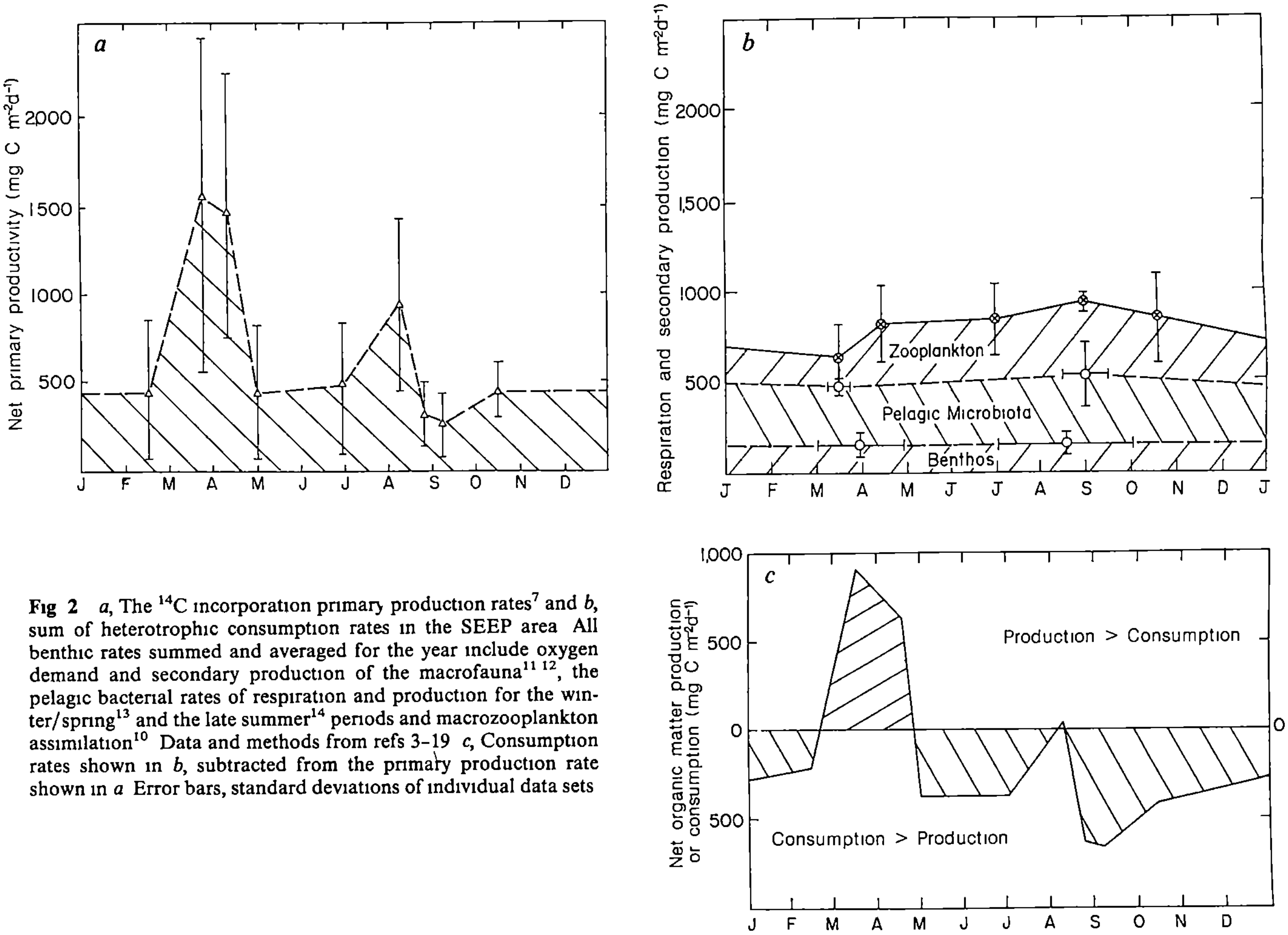


Fig 2 *a*, The ^{14}C incorporation primary production rates[7] and *b*, sum of heterotrophic consumption rates in the SEEP area All benthic rates summed and averaged for the year include oxygen demand and secondary production of the macrofauna[11 12], the pelagic bacterial rates of respiration and production for the winter/spring[13] and the late summer[14] periods and macrozooplankton assimilation[10] Data and methods from refs 3-19 *c*, Consumption rates shown in *b*, subtracted from the primary production rate shown in *a* Error bars, standard deviations of individual data sets

We have plotted the two separately over the course of a year (Fig 2*a*, *b*) The seasonal contrast is obvious on subtraction of integrated consumption from production (Fig 2*c*) This suggests that carbon fixation was predominant in the spring but that total remineralization and secondary production were the most significant processes during the rest of the year The results presented in Fig 2*c* provide no evidence to support the hypothesis that significant quantities of organic matter are available for export An annual manifestation of the intense heterotrophic activity following the bloom is the recurrent hypoxia experienced in the mid-Atlantic bight every year during stratified periods[19]

The budgeting approach undoubtedly suffers from some inaccuracies due to imprecision in each component The ^{14}C method may underestimate primary productivity (Fig 2*a*) if there is a loss of smaller photosynthetic picoplankton[20] Consumption (Fig 2*b*), on the other hand, does not include microzooplankton But as the microzooplankton consume the picoplankton as well as bacteria, the two omissions could to some degree cancel each other out[21] Sediment anaerobic respiration that is not incorporated into the measured total sediment oxygen demand (a distinct possibility if sulphide forms pyrite or is otherwise immobilized[12]) has also not been taken into account In spite of the limitations of our data set, the negative imbalance in the carbon budget (Fig 2*c*) cannot be considered justification for accepting the export hypothesis

Fluxes of organic carbon into sediment placed on the west line (Fig 1) also provided little evidence for accepting the hypothesis One mooring array of two traps was set for three different periods between 25 February and 7 April 1984, within continental slope water but underlying continental shelf water (locality W12 on Fig 1) Extensive particle export by sinking through or mixing across the shelf-slope front should be reflected in the traps, but the average rate of organic carbon flux to all traps (Table 1) was only 128 mg C m^{-2} d^{-1} ($\sigma = \pm 45$), or about 10% of the integrated euphotic zone primary productivity ($\sim$1 5 g C m^{-2} d^{-1}, $\sigma = \pm 0$ 76) at that time[7] This small fraction of the overlying net primary productivity was about what would be expected from direct vertical transport[22,23] and was below what we would expect if 90% of the net primary production during the spring bloom was exported to this depth[1,2] If only 50% of the net productivity was exported from a 100 km-wide continental shelf, and concentrated on a 30 km-wide continental slope, we would expect the organic carbon particle flux to be $\sim$1-2 g C m^{-2} d^{-1}, which is ten times greater than that sampled by the traps (Table 1)

Concurrent investigations of the transport of plant pigments and the characteristics of the sediments and biota of the continental slope support the above conclusions Although fluxes into sediment traps lower on the slope were greater near bottom than in the water column, the flux rates were far less than those we measured (Table 1) adjacent to the shelf-slope front[6] The moored fluorometers documented rapid sinking and extensive along-shore advection of biogenic particles Although some equivocal evidence suggested a net offshore transport may have occurred on the shelf, there was no evidence that appreciable material crossed the shelf-slope front[4] Cross-front exchanges of water between distinct shelf and slope water masses were very limited during the spring[24], although turbulent mixing may have been accentuated at the bottom by internal waves[25] An exchange of water of $\sim$20% across the front as it moves from Cape Cod to Cape Hatteras, based on salt and nitrate budgets[7], could be accounted for by this bottom mixed-layer exchange[25]

The logic behind extensive organic carbon export has been questioned before[26,27], but our data graphically illustrate that seasonal coupling and pelagic microbial activity[28] make the idea untenable Our results do not support the suggestion that an appreciable fraction of the global missing carbon is on continental slopes[1], but they do not exclude the possibility that most of the carbon on continental slopes was due to shelf export, albeit at a relatively slow rate Although our rejection of the hypothesis applies to the north-west Atlantic SEEP area in particular, it is important to ask how our conclusions might apply to other shelf ecosystems At higher latitudes, for example, colder water may depress microheterotrophic activity[29] The export and fate of carbon from narrow shelves where primary production is enhanced by external sources of 'new' nutrients[2,30,31] remains to be assessed

This work was part of continuing studies of coastal zone ecosystems by the Office of Health and Environmental Research, US Department of Energy (contract DE-ACO2-76CH00016), and the Northeast Fisheries Center, NMFS/NOAA We thank our collaborators in the SEEP experiments, as well as the officers and crew of the RV *Oceanus*

Received 22 May, accepted 25 September 1986

1 Walsh, J J, Rowe, G T, Iverson, R L & McRoy, C P *Nature* **291**, 196-201 (1981)
2 Malone, T, Hopkins, T S, Falkowski, P C & Whitledge, T *Cont Shelf Res* **1**, 305-337 (1983)
3 Whitledge, T E & Wirick, C D *Deep Sea Res* **30**, 29-309 (1983)
4 Walsh, J J, Wirick, C, Pietrafesa, L & Whitledge, T & Hoge, F *Cont Shelf Res* (in the press)
5 Rowe, G & Gardner, W *J mar Res* **37**, 581-600 (1979)
6 Biscaye, P & Anderson, R. *Cont Shelf Res* (submitted)
7 Falkowski, P, Rowe, G, Smith, S, Whitledge, T & Flagg, C *Cont Shelf Res* (in the press)
8 Dagg, M J & Grill, D *Limnol Oceanogr* **25**, 597-609 (1980)
9 Dagg, M J & Turner, J F *Can J Fish aquat Sci* **39**, 979-990 (1982)
10 Smith, S & Lane, P V Z *Cont Shelf Res* (in the press)
11 Florek, R & Rowe, G *Int Rev ges Hydrobiol* **68**, 73-112 (1983)
12 Rowe, G T *et al. Cont Shelf Res* (in the press)
13 Ducklow, H & Kirchman, D L *J Plank. Res* **5**, 333-355 (1983)
14 Harrison, W G, Douglas, D, Falkowski, P, Rowe, G & Vidal, J *J Plank. Res* **5**, 539-556 (1983)
15 Warwick, R. M, George, C & Davies, J R. *Est Coast mar Sci.* **7**, 2-5 (1978)
16 Banse, K & Mosher, S *Ecol. Monogr* **50**, 355-379 (1980)
17 Wigley, R L & Theroux, R. B *Prof Pap U S geol. Surv* **529-N** (1981)
18 Collie, J S *Mar Ecol. Prog Ser* **22**, 229-238 (1985)
19 Falkowski, P G, Hopkins, T S & Walsh, J J *J mar Res* **38**, 479-506 (1980)
20 Glover, H E, Campbell, L & Prezlin, B B *Mar Biol.* **91**, 193-203 (1986)
21 Goldman, J C & Caron, P A *Deep Sea Res* **32**, 899-915 (1985)
22 Suess, E *Nature* **288**, 260-263 (1980)
23 Lorenzen, C J, Welschmeyer, N & Copping, A E *Deep Sea Res* **30**, 639-643 (1983)
24 Houghton, R W, Aikman, F & Ou, H W *Cont Shelf Res* (in the press)
25 Flagg, C N *Cont Shelf Res* (in the press)
26 Peng, T-H & Broecker, W *J geophys Res* **89**, 8170--8180 (1984)
27 Pace, M L, Glasser, J & Pomeroy, L R. *Mar Biol.* **82**, 47-63 (1984)
28 Ducklow, H W, Purdie, D A, Williams, P LeB & Davies, J M *Science* **232**, 865-867 (1986)
29 Pomeroy, L R & Deibel, D *Science* **233**, 359-361 (1986)
30 Walsh, J J *Nature* **290**, 300-304 (1981)
31 Rowe, G & Howarth, R *Deep Sea Res* **32**, 43-55 (1985)

Identification of novel widely distributed sedimentary acyclic sesterterpenoids

J. N. Robson & S. J. Rowland

Department of Environmental Sciences, Plymouth Polytechnic,
Drake Circus, Plymouth, Devon PL4 8AA, UK

Isoprenoids, the general class of natural products biosynthesized from isoprene units, are abundant in nature. Many isoprenoids are also found in sediments, sedimentary rocks and crude oils where they have proved useful as 'chemical fossils' of biological activity and as indicators of geothermal stress[1]. However, acyclic isoprenoids with 25 carbon atoms—so-called sesterterpenoids—have been reported only rarely in the biosphere and the geosphere. Indeed, the only acyclic sesterterpenoids of confirmed structure reported in sediments are of pentamethyleicosane structural isomers (Fig. 1, carbon skeleton I). For example, 2,6,10,14,18-pentamethyleicosane has been proposed as a marker for halophilic bacteria and the 2,6,10,15,19-isomer (I) as an indicator of methanogens[2]. We have now identified a group of novel sesterterpenoids that are major components of the hydrocarbons of recent freshwater, marine and hypersaline sediments from many different parts of the globe. Because they have unusual structures, these sesterterpenes show promise as a new class of biological marker compounds.

Over the past 10 years many workers (Table 1) have reported the presence, in recent sediments, of a group of C_{25} hydrocarbons of unknown structure The compounds are abundant and are very widely distributed, for example, they are the major hydrocarbons present in sediments from a New England salt marsh, Narragansett Bay (USA)[3], North-east Gulf of Mexico[4], the Peru upwelling region[5], Shark Bay, Western Australia[6] and the Ebro River, Spain[7] The same compounds have been detected in aquatic filtering organisms and in sediment traps suspended in the water column[7-10] and recently a C_{25} diene has been detected in field specimens of the green alga, *Enteromorpha*[11] Numerous structures have been proposed for these C_{25} compounds (see Fig 1, skeletons II-V)[7,11,12] although no structure has been confirmed The electron impact mass spectrum and gas chromatographic Kovats Index (KI) of the parent C_{25} alkane[12,13] and ozonolysis of a C_{25} monoene[6] suggest structure II as the most likely skeleton This is supported by synthesis of a co-occurring

Table 1 Occurrence of alkanes and alkenes with skeletons II, VII and VIII in sediments*

Locations	Age/environment	Carbon skeleton II	VII	VIII	Refs
Kiel Bight, FRG	Recent/marine	√		√	17, 18
Scotian Shelf, Nova Scotia	Recent/marine	√			19
Shark Bay, Western Australia	Recent/marine/hypersaline	√	√		6
Alaskan Outer Continental Shelf	Recent/marine	√		√	20
Peru Continental Shelf	Recent/marine	√			5, 21
Alfacs Bay, Ebro River, Spain	Recent/marine	√	√		7, 22
Dabob Bay, Washington State, USA	Recent/land-locked marine	√		√	9
Northeast Gulf of Mexico	Recent/marine	√	√		4
Southern California, USA	Recent/marine	√			10
Rhode Island Sound, USA	Recent/estuarine	√			8
Buzzards Bay, Massachusetts, USA	Recent/estuarine	√			23
Narragansett Bay, Rhode Island, USA	Recent/estuarine	√		√	3, 15
Puget Sound, Washington State, USA	Recent/land-locked marine	√	√	√	12, 13
Sandyhaven, Dyfed, Wales	Recent/intertidal	√	√		11
Grasmere, Cumbria, UK	Recent/freshwater lake		√		14
Camloch, Scotland	Recent/freshwater lake		√		24
Loch Clair, Scotland	Recent/freshwater lake		√		24
Upton Broad, Norfolk, UK	Recent/freshwater lake		√		24
Rostherne Mere, Cheshire, UK	Recent/freshwater lake	√	√		11
Cariaco Trench, Venezuela	Pleistocene/marine		√		25
Sabon-Gida, Nigeria	Pleistocene/lacustrine	√	√		25
Loch Ewe, Scotland	Recent/marine loch		√		25
Esthwaite Water, Cumbria, UK	Recent/freshwater lake		√		26
Pettaquamscutt River, Rhode Island, USA	Recent/estuarine	√		√	27
Sullom Voe, Shetland Islands, UK	Recent/intertidal	√	√		28

* The same compounds have been found in marine filtering organisms[8], an alga[11] and certain crude oils[25]

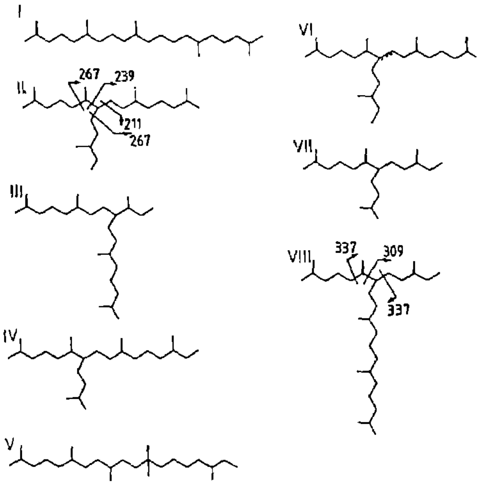


Fig 1 Isoprenoids, carbon skeletons I-VIII

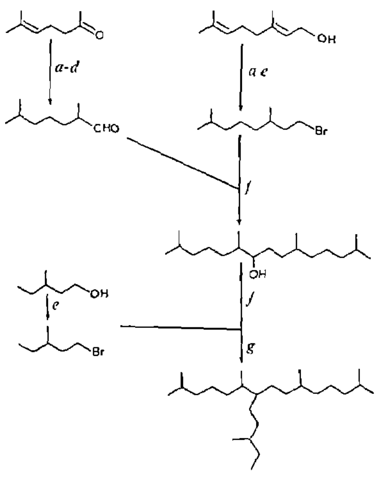


Fig 2 Synthetic scheme for alkane II and alkenes VI All products characterized by ^{1}H NMR mass and infrared spectroscopy The second Grignard reaction (f) was repeated in improved yield (82%) by use of CeCl$_3$/Mg (ref 29) a, H$_2$,PtO$_2$,hexane, b, BrCH$_2$Br,Mg, c, B$_2$H$_6$,THF,H$_2$O$_2$, OH d, (COCl)$_2$,DMSO, e, HBr/H$_2$SO$_4$, f, Mg,THF, g, POCl$_3$,Pyr

C_{20} alkane (assumed to be pseudohomologous with the C_{25} alkane)[14] Synthesis of V has shown this structure to be incorrect[12] and attempts to isolate the C_{25} alkenes by argentatious thin layer chromatography in quantities large enough for spectral characterization have been unsuccessful[15]

A mixture of five monoenes related to alkane II was synthesized in good yield from readily available starting materials by the scheme outlined in Fig 2 Catalytic hydrogenation afforded the alkane The electron impact mass spectrum of the synthetic alkane II is shown in Fig 3 The spectrum, which is virtually identical to that of the sedimentary alkane[12] contains small but characteristic ion doublets at *m/z* 210,211, 238,239 and 266, 267 assigned to fragmentation about the C7-C8, C6-C7 and C5-C6 (and C7-C1') bonds (structure II) The KI value for both synthetic and sedimentary alkane is 2,109 (DB-1, 40-290 °C at 4 °C min^{-1}) These data confirm that the C_{25} alkenes and alkane identified in many previous studies (Table 1 and ref 11) possess the 2,6,10,14-tetramethyl-7-(3-methylpentyl) pentadecane skeleton (structure II) Mass spectra of the synthetic monoenes (Fig 1, structure VI) were characterized by even mass ions at *m/z* 154,196,210 (weak), 224,266,280 and 350(M +) but none was identical to the spectra of any of the sedimentary monoenes This is further indirect evidence that the sedimentary alkenes may indeed contain a methylene double bond as suggested previously from ozonolysis[6] and nuclear magnetic resonance (NMR)[8] In sediments from Puget Sound[12] and Dabob Bay[9], Washington State (USA), the C_{25} hydrocarbons co-occurred with the C_{20} alkane (now known to be structure VII in Fig 1[11]) and a C_{30} alkene of unknown structure but which from retention index measurements was considered to be pseudohomologous with both structures VII and II when hydrogenated Given the structural relationship between VII and II demonstrated in the present study, it is reasonable to speculate that the C_{30} alkene is VIII (Fig 1) This is consistent with the mass spectrum in that fragmentation about the C6-C7, C5-C6

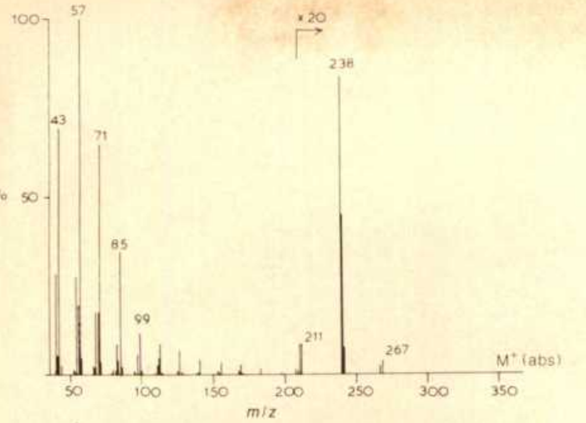


Fig. 3 Electron impact mass spectrum of synthetic II. Gas chromatography mass spectrometry conditions: Kratos MS25, 30 m × 0.3 mm CPSil 5 fused silica column, 40–300 °C at 4 °C min^{-1}. He carrier, 40 eV.

and C7–C1 bonds would produce the ions observed at m/z 308,309; 336,337 (structure VIII). Thus, the abundant sesterterpenes appear to be the major and most commonly occurring components of a new class of terpenoids comprising $C_{20,25,30}$ and possibly other[6], hydrocarbons.

These highly branched hydrocarbons have already proved useful as environmental indicators. For instance, variations in their abundance in sediments have been successfully used to designate distinct chemogeographic regions. Dunlop and Jefferies[6] were able to classify oceanic and hypersaline sedimentary basins in Western Australia on this basis and Requego and Quinn[15] were able to show that the geographical variations of some of the alkenes reflected the distribution of marine organic matter in the estuary of Narragansett Bay, USA. Now that the structures of the C_{20} and C_{25} hydrocarbons have been established and one possible source organism proposed[11], future studies may be based on a firmer understanding. Further synthesis and incubation of synthetic ^{14}C-labelled analogues should help to elucidate the longer term sedimentary fate of these new biological markers.

Finally, the unusual structures of these molecules suggest that they may have a specific role in the lipid biochemistry of the source organisms. This has proved to be the case for other unusual isoprenoid lipids. For instance biphytanyl isoprenoid ethers act as rigidifiers in the cell membranes of archaebacteria[16] and are therefore of interest to biologists, biochemists and geochemists. Further study of the present multi-branched lipids may benefit from a similar multidisciplinary approach.

Received 1 September; accepted 3 October 1986.

1. Tissot, B. P. & Welte, D. H. *Petroleum Formation and Occurrence* (Springer, Berlin, 1984).
2. Brassell, S. C., Wardroper, A. M. K., Thompson, I. D., Maxwell, J. R. & Eglinton, G. *Nature* **290,** 693–696 (1981).
3. Requejo, A. G. & Quinn, J. G. *Estuar. Coast. Shelf Sci.* **20,** 281–297 (1985).
4. Gearing, P., Gearing, J. N., Lytle, T. F. & Lytle, J. S. *Geochim. cosmochim. Acta* **40,** 1005–1017 (1976).
5. Volkman, J. K., Farrington, J. W., Gagosian, R. B. & Wakeham, S. G. in *Advances in Organic Geochemistry, 1981* (eds Bjoroy, M. *et al.*) 228–240 (Wiley, New York, 1983).
6. Dunlop, R. W. & Jefferies, P. R. *Org. Geochem.* **8,** 313–320 (1985).
7. Albaiges, J. *et al. Org. Geochem.* **6,** 237–248 (1984).
8. Boehm, P. D. & Quinn, J. G. *Estuar. Coast. mar. Sci.* **6,** 471–494 (1978).
9. Prahl, F. G., Bennett, J. T. & Carpenter, R. *Geochim. cosmochim. Acta* **44,** 1967–1976 (1980).
10. Crisp, P. T., Brenner, S., Venkateson, M. I., Ruth, E. & Kaplan, I. R. *Geochim. cosmochim. Acta* **43,** 1791–1801 (1979).
11. Rowland, S. J., Yon, D. A., Lewis, C. A. & Maxwell, J. R. *Org. Geochem.* **8,** 207–213 (1985).
12. Barrick, R. C. & Hedges, J. I. *Geochim. cosmochim. Acta* **45,** 381–392 (1981).
13. Barrick, R. C., Hedges, J. I. & Peterson, M. L. *Geochim. cosmochim. Acta* **44,** 1349–1362 (1980).
14. Yon, D. A., Ryback, G. & Maxwell, J. R. *Tetrahedron Lett.* **23,** 2143–2146 (1982).
15. Requejo, A. G. & Quinn, J. G. *Geochim. cosmochim. Acta* **47,** 1075–1090 (1983).
16. De Rosa, M., De Rosa, S., Gambacorta, A., Minale, L. & Bu'Lock, J. D. *Phytochemistry* **16,** 1961–1965 (1977).
17. Osterroht, C. & Petrick, G. *Mar. Chem.* **11,** 55–70 (1982).
18. Ehrhardt, M. & Heinemann, J. *Environ. Pollut.* **9,** 263–282 (1975).
19. Keizer, P. D., Dale, J. & Gordon, D. C. *Geochim. cosmochim. Acta* **42,** 165–172 (1978).
20. Venkatesan, M. I. & Kaplan, I. R. *Geochim. cosmochim. Acta* **46,** 2135–2149 (1982).
21. Smith, D. J., Eglinton, G. & Morris, R. J. *Geochim. cosmochim. Acta* **47,** 2225–2232 (1983).
22. Bayona, J. M. *et al. J. High Res. Chromatography chromatogr. Commun.* 605–611 (1983).
23. Farrington, J. W., Frew, N. M., Gschwend, P. M. & Tripp, B. W. *Estuar. Coastal Mar. Sci.* **5,** 793–808 (1977).
24. Cranwell, P. A. *Prog. Lipid Res.* **21,** 271–308 (1983).
25. Yon, D. A. thesis, Univ. Bristol (1981).
26. Brooks, P. W. *et al. Chem. Geol.* **20,** 189–204 (1977).
27. Requejo, A. G., Quinn, J. G., Gearing, J. N. & Gearing, P. J. *Org. Geochem.* **7,** 1–10 (1984).
28. Jones, D. M. thesis, Univ. Newcastle-on-Tyne (1986).
29. Imamoto, T., Takiyama, N. & Nakamura, K. *Tetrahedron Lett.* **26,** 4763–4766 (1985).

Ostrich ancestors found in the Northern Hemisphere suggest new hypothesis of ratite origins

Peter Houde

Division of Birds, National Museum of Natural History,
Smithsonian Institution, Washington, D.C. 20560, USA

Modern ratites (ostriches, rheas, cassowaries, emus, and kiwis) are flightless birds which have a palatal structure termed 'palaeognathous'[1] and are found on daughter-landmasses of the Mesozoic supercontinent Gondwanaland. It has been suggested[2–4] that a single flightless ancestor, widely distributed in Gondwanaland, gave rise to the various lineages of ratite birds. The temporal calibration of the DNA molecular clock is primarily based on the divergence of ratites, and depends on the validity of this hypothesis. Newly studied fossils suggest that the ancestors of ostriches are instead among a group of North American and European birds, the '*Lithornis*-cohort', that had the potential of flight and from which the kiwis may have arisen separately.

The *Lithornis*-cohort of fossils is a group of hen-sized, palaeognathous carinates which had the potential of flight and are found in Palaeocene and Eocene deposits in North America and Europe[5]. The group includes *Lithornis*, *Paracathartes* and a genus I will refer to here as the 'Green River palaeognath' ('GRP'). They are the closest relatives of the ratite birds known to have had the capacity for flight, closer than the palaeo-

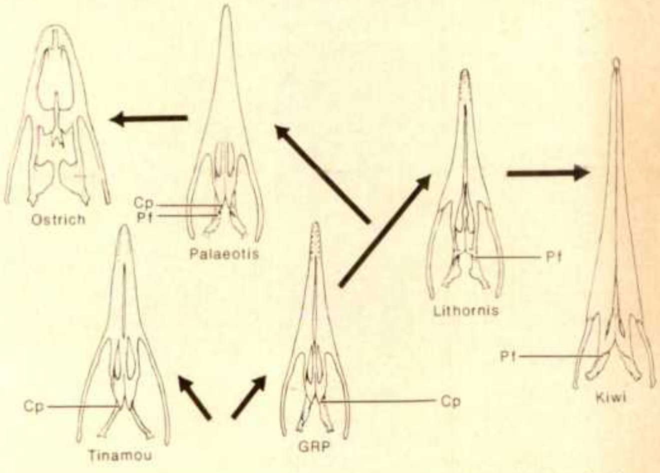


Fig. 1 Grades of evolution in the palatal complex of bones in the fossil and extant palaeognathous birds: tinamou (*Tinamus*, drawn from Smithsonian specimen USNM 345738); 'Green River palaeognath' (drawn from USNM 336103); *Lithornis* (drawn from USNM 391983; agrees with *Paracathartes*, USNM 391984); *Palaeotis* (drawn from HLMD Me 7530); kiwi (drawn from USNM 19025) and ostrich (drawn from USNM 224852), not to scale. The branching pattern does not necessarily represent the correct phylogenetic relationships of the fossil taxa but the monophyly of the ostrich and kiwi and the sister group relationship of tinamous is supported by molecular data[4]. Cp, caudal process of the palatine; Pf, pterygoid fossa.

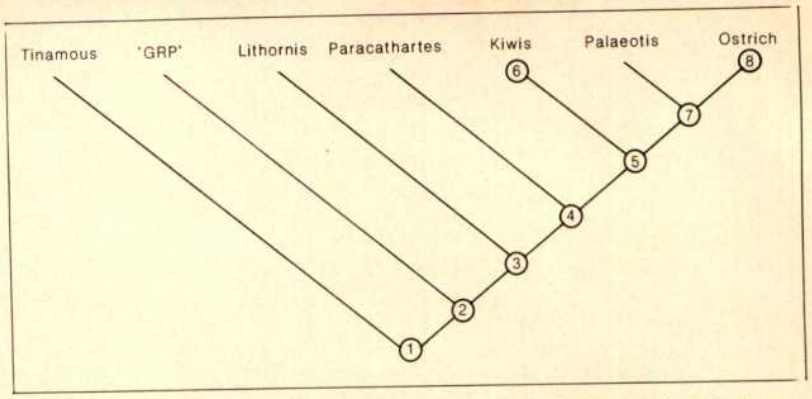


Fig. 2 Cladogram of some taxa of palaeognathous birds of unequal rank. Ratites without direct bearing to this discussion are omitted. Character suites are presumed to be apomorphous for all clades above 1. The caudal process of palatine is convergently absent in *Lithornis*, *Paracathartes* and Kiwis (not shown in cladogram). Character suites are: (1) palate palaeognathous, caudal process of palatine bone present, pterygoid narrow, vomer articulates with premaxillae, rhamphothecal grooves present, internarial septum small, splenial extends near to or forms part of mandibular symphysis, bill narrow, cervical vertebrae stout, ilioischiadic foramen open-ended, femur gracile, primary osteons of tibiotarsus predominantly oriented longitudinally, tibiotarsus with large lateral cnemial crest and well differentiated medial and lateral condyles, tarsometatarsus with pronounced, spherical intercotylar eminence, four toes present; (2) internarial septum large, pterygoid broad, quadrate fits into V-shaped notch in mandible, buttressed from behind by retroarticular process of mandible, scapula with styloid acromion, tarsometatarsus with large interosseous foramen, deeply concave anterior face and deeply excavated origin of extensor digiti IV brevis muscle; (3) pterygoid fossa present, vomer does not articulate with premaxillae; (4) primary osteons of tibiotarsus predominantly oriented transversely; (5) scapulocoracoid fused and sternum unkeeled, humerus long and spindly, with pronounced internal tuberosity and without pneumatic foramen, antebrachium short; (6) bill elongate, eyes reduced, tarsometatarsus without concave anterior face and excavated origin of extensor digiti IV brevis muscle, interosseous foramen small; (7) scapulocoracoid with large glenoid, acromion and acrocoracoid process, carpometacarpus with accessory foramina between each of the metacarpal bones, ilioischiadic foramen elongate, femur short, tarsometatarsus long, hallux absent (condition of quadrate mandible articulation unknown); and (8) pterygoid fossa absent, caudal process of palatine bone absent, V-shaped notch in mandible for quadrate absent, bill broad, cervical vertebrae elongate, femur short and robust, tibiotarsus with pronounced anterior cnemial crest and undifferentiated medial and lateral condyles, tarsometatarsus without spherical intercotylar eminence, concave anterior face and excavated origin of extensor digiti IV brevis muscle, interosseous foramen small and only two trochleae present.

gnathous tinamous (Tinamidae), the extant sister group of ratites[1,2,4,6]. *Palaeotis weigelti* was a crane-sized, flightless palaeognathous ostrich from the Middle Eocene of Germany, whose ancestry may be traced to the *Lithornis*-cohort. *Lithornis*[7], *Paracathartes*[8] and *Palaeotis*[9] were originally described from isolated, fragmentary remains and were mistaken for neognathous birds. Newly available, essentially complete skeletons that include the palatal bones (Fig. 1) now show unequivocally that these were palaeognathous birds. *Palaeotis* should not be confused with the incomplete fossils comprising the Asian family Ergilornithidae that have been suggested as ancestors of ostriches[10,11]. Ergilornithids were not palaeognathous birds.

The monophyletic origin of the *Lithornis*-cohort can be shown on the basis of shared derived characters (Fig. 2, character suite 2), all of which are absent in tinamous. Cranial (Fig. 1) and histological characters (Fig. 3) that characterize the living ratites evolved step by step in the phylogeny of the *Lithornis*-cohort. The large pterygoid fossa with no palatine process first appeared in clade 3, and the longitudinal pattern of bone osteons was replaced by the transverse pattern in clade 4 (Fig. 2). A higher-taxonomic group, comprised of the *Lithornis*-cohort, would thus be paraphyletic. That is, it does not include all of its evolutionary derivatives (the ratites), which have secondarily lost some of the distinguishing flight-related characters of the *Lithornis*-cohort.

An important element in the elucidation of ratite phylogeny is the fossil *Palaeotis*. Two skeletons of *Palaeotis weigelti* (specimen GM 4362, Geiseltalmuseum, Martin-Luther University, Halle a. Saale, GDR and specimen HLMD Me 7530, Hessisches Landesmuseum, Darmstadt, FRG) reveal all the characters that together define ratite birds,[1,2,6]. Derived characters that further distinguish *Palaeotis* as an ostrich (family Struthionidae) are

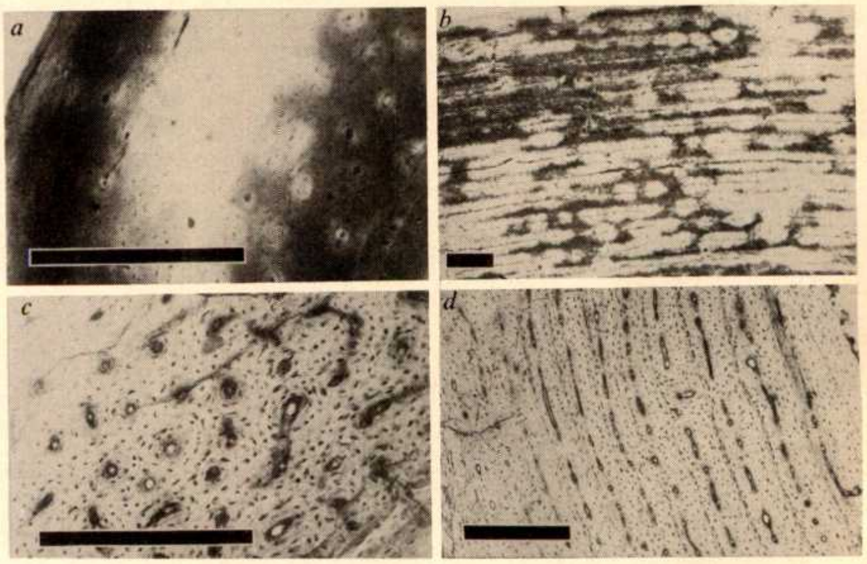


Fig. 3 Bright field photomicrographs of transverse lapidary thin sections through the same level (distal metaphysis) of *a*, tinamou (*Rhynchotus*, Princeton University PU 728); *b*, ratite (*Rhea*, specimen unnumbered); *c*, *Lithornis* (USNM 290554); and *d*, *Paracathartes* (USNM 361407), showing histological character states used in Fig. 2. Vascular canals appear as thin dark lines, spots, or circles on a light background of inorganic bone. Large dark regions are more opaque regions of bone, in which vascular canals may appear as light spots. Scale, 0.3 mm. The patterns formed by primary osteons are characteristically different in the leg bones of palaeognathous birds (tinamous and ratites) and neognathous birds (all other birds)[16,17]. The primary osteons in the tibiotarsus of *Lithornis* are predominantly parallel to the longitudinal axis of the bone, as in the tibiotarsus of the palaeognathous tinamous but unlike ratites and neognathous birds. Primary osteons in the tibiotarsus of *Paracathartes* are predominantly arranged in concentric rings that lie in the transverse plane of the bone, as in the tibiotarsus of ratites but unlike the tinamous and neognathous birds. Although the patterns of primary osteons could be evolved separately by unrelated birds, no correlation has been found between the pattern of primary osteons and flightlessness, ontogeny, or overall size among neognathous birds[18].

listed (Fig. 2, legend, character suite 7). Characters that are derived at the level of the *Lithornis*-cohort (Fig. 2, character suite 2) and primitively retained by *Palaeotis*, but not found in other birds, show that ostriches evolved from this group. Only *Palaeotis* and members of the *Lithornis*-cohort have palates with a caudal process of the palatine and a pterygoid fossa (Fig. 1), although both characters are not found together in *Lithornis*, *Paracathartes*, or the 'GRP'. *Palaeotis* forms an important link between the *Lithornis*-cohort and ostriches because all these characters are absent in large ratites, having been secondarily lost by all but kiwis.

Thus, ostriches were not only present in the Northern Hemisphere, but they may have evolved there long after the breakup of Gondwanaland. If so, then ostriches emigrated from Europe to Africa some time during the early or middle Tertiary, a dispersal route for which there are independent data for other taxa[12]. The hypothesis of ratite origins in Gondwanaland relies on the same route of dispersal, but in the opposite direction, to account for the presence of the recently extinct ostriches of Asia[2,13]. The absence of the *Lithornis*-cohort in early Tertiary deposits of the Southern Hemisphere does not conclusively prove that these birds lived only in the Northern Hemisphere at that time, but their absence there is conspicuous since *Lithornis* has been collected in relative abundance in the Northern Hemisphere for more than two centuries.

The vicariance biogeography hypothesis of ratite distribution requires that populations of flightless ratites were isolated as the result of geotectonic events and that the birds' own powers of dispersal were minimal. However, this hypothesis does not explain the occurrence of kiwis (Apterygidae) in New Zealand. Sibley and Ahlquist[4] could not reconcile the early (Cretaceous) divergence of the New Zealand plate from Gondwanaland with the relatively late (Eocene) divergence of kiwis from other ratite birds, based on DNA hybridization data. The fossils discussed here suggest that kiwis could have evolved directly from a palaeognath that flew to New Zealand and lost its power of flight independently of other ratite birds. Of the ratites, only kiwis retain primitive characters that are uniquely derived in the *Lithornis*-cohort (Fig. 1; Fig. 2, character suites 2 and 3), suggesting a relatively close relationship.

The far reaching implications of this study lie with the temporal calibration of the DNA hybridization molecular clock. Sibley and Ahlquist[4] postulated that the divergence of ostrich DNA from rhea DNA was initiated by, and therefore could be dated by, the spreading of the Atlantic seafloor. They have since[14] applied this calibration to a variety of other avian and mammalian taxa. With the simplicity of the vicariance biogeography hypothesis of ratite origins now challenged, one cannot accept this part of the calibration of the DNA molecular clock without at least some degree of scepticism. I have elsewhere[15] discussed problems with the four other data used to calibrate the DNA molecular clock.

I thank A. J. Charig for his comments on this manuscript and D. Baird, J. Becker, D. P. Domning, A. Elzanowski, P. D. Gingerich, H. Haubold, A. E. Houde, W. v. Koenigswald, S. L. Olson, A. de Ricqlés, G. Stanton and R. L. Zusi for their assistance. This work was supported by the A. Wetmore Memorial Fund of the American Ornithologists Union, A. Wetmore Memorial Fund of the Smithsonian Institution, F. M. Chapman Memorial Fund of the American Museum of Natural History, Museum of Paleontology of the University of Michigan, National Science Foundation Dissertation Grant (BSR 8313209), Sigma Xi, Smithsonian Predoctoral Fellowship and National Institutes General Medical Sciences Grant (GM07803) to L. V. Leak of Howard University.

Received 23 January; accepted 18 September 1986.

1. Bock, W. J. *Proc. XIII Int. Congr. Orn.* 39-54 (1963).
2. Cracraft, J. *J. Zool.* **169,** 455-545 (1973).
3. Prager, E. M., Wilson, A. C., Osuga, D. T. & Feeney, R. E. *J. molec. Evol.* **8,** 283-294 (1976).
4. Sibley, C. G. & Ahlquist, J. E. *Evolution Today, Proc. Second int. Congr. Syst. evol. Biol.* (eds Scudder, G. G. E. & Reveal, J. L.) 301-335 (Carnegie-Mellon University, Pittsburgh).
5. Houde, P. & Olson, S. L. *Science* **214,** 1236-1237 (1981).
6. Parkes, K. C. & Clark, G. A. Jr *Condor* **68,** 459-471 (1966).
7. Owen, R. *Proc. Geol. Soc. Lond.* **3,** 162-166 (1840).
8. Harrison, C. J. O. *Tertiary Res. spec. Pap.* **5,** 7-10 (1979).
9. Lambrecht, K. *Jb. hallesch. Verb. Erforsch. Mitteldt. Bodenschatze* **7,** 1-11 (1928).
10. Feduccia, A. *The Age of Birds* (Harvard University Press, Cambridge, 1980).
11. Olson, S. L. in *Avian Biology* Vol. 8 (eds Farner, D. S., King, J. R. & Parkes, K. C.) 80-238 (Academic, New York, 1985).
12. Van Couvering, J. A. & Van Couvering, J. A. H. *Proc. VI Congr. Brataslava, Reg. Com. Mediterranean Neogene Strat.*, 363-367 (Brataslava, Czechoslovakia, 1975).
13. Cracraft, J. *Ibis* **116,** 494-521 (1974).
14. Sibley, C. G. & Ahlquist, J. E. *J. molec. Evol.* **20,** 2-15 (1984).
15. Houde, P. *Auk* (in the press).
16. Amprino, R. & Godina, G. *Anat. Anz.* **95,** 191-214 (1944).
17. Zavattari, E. & Cellini, I. *Monitore zool. ital.* **64,** 189-200 (1956).
18. Houde, P. *Auk* (in the press).

Phospholipid vesicles as a model system for biomineralization

Stephen Mann*, John P. Hannington*
& R. J. P. Williams†

* School of Chemistry, University of Bath, Bath BA2 7AY, UK
† Inorganic Chemistry Laboratory, South Parks Road,
Oxford OX1 3QR, UK

The tailoring of inorganic minerals such as iron oxides for functional use in biological systems for iron storage[1], structural support[2] and magnetoreception[3] involves biological regulation of crystal structure, particle size, morphology and crystallographic organization. The encapsulation of crystallochemical reactions within enclosed biological microvolumes enables control to be exerted over: (1) the chemical regulation, by passive or facilitated ion-transport, of localised supersaturation levels (2) the stereochemical requirements for ion-binding, redox and nucleation events at the organic matrix interface and (3) the spatial organization of crystal growth and morphology[4,5]. Matrix-mediated growth of inorganic materials has not been systematically investigated *in vitro* even though it may have applications in crystal engineering and materials science[6]. We have used phospholipid unilamellar vesicles of ~300 Å diameter to study membrane-mediated processes of iron oxide crystal growth. Intravesicular deposits differ in structure, morphology and size from precipitates formed from reactions in bulk aqueous solution. Mediating factors include vesicle shape and dimension, diffusion-limited processes of ion-transport and ion-binding at the curved lipid headgroup surface. These results help elucidate biomineralization and have technological relevance to the controlled synthesis of monodisperse sols with catalytic and magnetic properties.

Figure 1 illustrates the general reaction scheme adopted in our work. Unilamellar phospholipid vesicles with diameters of ~300 Å were prepared from phosphatidylcholine (17 mM) sus-

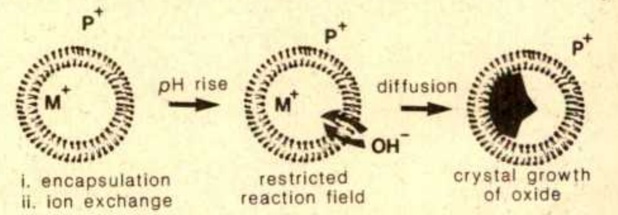


Fig. 1 Reaction scheme for membrane-mediated crystal growth of metal oxides in phospholipid vesicles. Cations (M^+) are encapsulated by sonication and replaced in the external phase by inert cations (P^+) by ion-exchange chromotography. This procedure results in organised restricted reaction fields for crystal growth. Increases in extravesicular *p*H result in OH^- influx and the nucleation and growth of metal oxides within the intravesicular space.

Table 1 Electron diffraction data for intravesicular precipitates

d spacing (Å) and associated Miller indices Intravesicular solution (see text)		
[Fe(III)]	[Fe(II)]	[Fe(II)/Fe(III)]
2.6 (021)	2.93 (220)	2.5 (110)
2.08 (140)	2.52 (311)	1.95 (113)
1.8 (211)	2.06 (400)	1.49 (106)
1.49 (250)	1.75 (422)	
1.32 (321)	1.44 (440)	
	1.24 (622)	
	1.05 (800)	
	1.00 (822)	
	0.92 (840)	

Lines of strong intensity are underlined. Miller indices are shown in parentheses.

pensions at 4 °C in the presence of Fe(II) and/or Fe(III) solutions. Three preparative routes were investigated: encapsulation in the presence of (1) 100 mM $FeCl_3$ at *p*H 2.0, (2) 100 mM $FeCl_2 4H_2O$ at *p*H 4.0, (3) 50 mM $FeCl_2$ and 50 mM $Fe(NO_3)_3$ solution at *p*H 2.0. In each experiment, metal ions external to the vesicles were removed by ion-exchange chromatography (Na-form). Precipitation of Fe(II) and/or Fe(III) ions in the restricted reaction field delineated by the vesicle membrane was then induced by addition of aqueous NaOH (final *p*H, 12.0) to the extravesicular solution. A pale yellow colouration was noted in all the solutions after 30 min and it increased in intensity over several days. No precipitation was observed after bench centrifugation. By contrast, in similar precipitation reactions in the absence of vesicles, rapid precipitation of iron oxides was apparent. Unstained samples were air dried onto electron microscope grids and studied by transmission electron microscopy, selected area electron diffraction and X-ray microanalysis.

Electron micrographs of iron-containing vesicles before *p*H increase and precipitation showed that the intravesicular iron was not present as a simple salt solution but that extensive adsorption via ion-binding at the choline headgroups had taken place (Fig. 2*a*). Precipitation of the intravesicular iron after NaOH addition resulted in the formation of discrete, finely divided particles with mean diameters 35–40 Å (Fig. 2*b*). X-ray microprobe analysis of vesicles before and after precipitation showed the presence of Fe, P and Cl (trace). None of these elements was detected in background analyses of the electron microscope grid. Table 1 shows the results obtained from electron diffraction identification of these intravesicular deposits. Intravesicular Fe(III) solutions crystallized as poorly-ordered spherulitic geothite (α-FeOOH) whereas extended aggregates of ferrihydrite were formed in bulk solution. Precipitation of intravesicular Fe(II) solutions eventually produced discrete 30–50 Å diameter magnetite (Fe_3O_4) crystallites (Fig. 2*b*) whereas bulk oxidation produced a mixture of acicular lepidocrocite (γ-FeOOH) and goethite. Oxidation and precipitation of initial intravesicular Fe(II)/Fe(III) solutions formed poorly-ordered ferrihydrite in contrast with the crystallization of 100–500 Å magnetite particles in the absence of vesicles.

These results can be rationalized in terms of the control by the vesicle membrane on the rate of OH^- diffusion into the intravesicular space. We have shown elsewhere[7] that the influx of hydroxyl ions into phospholipid vesicles is critically related to the efflux of counter-anions from the intravesicular volume. The net rate of OH^- diffusion and hence the rate of iron oxide formation will be dependent on the efflux rate of the slower-moving chloride anion[8]. The rate of intravesicular Fe(II) oxidation, which is dependent on the rate of increase of intravesicular *p*H, has been followed spectrophotometrically by measurement of the Fe(III) absorption band at 420 nm. Whereas bulk oxidation at neutral or alkaline *p*H is rapid and follows overall third-order kinetics at constant $[O_2]$ (ref. 9), intravesicular Fe(II) oxidation shows overall first-order behaviour ($k = 3.7 \times 10^{-7}\ s^{-1}$) (Fig. 3) indicative of a diffusion-controlled mechanism.

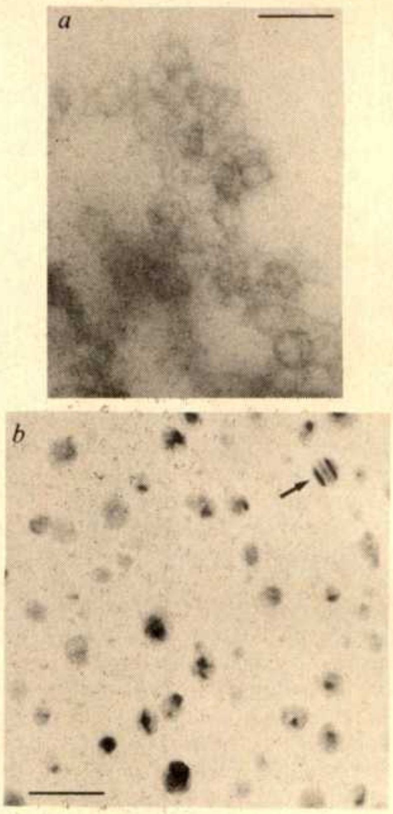


Fig. 2 *a*, Transmission electron microscope (TEM) image of phospholipid vesicles containing Fe(III) ions at *p*H 2.0. The binding of Fe(III) at the charged headgroups results in the *in situ* staining of the inner membrane surface. Bar, 60 nm. *b*, High magnification TEM image of intravesicular magnetite particles. Some particles show electron dense banding (arrow) which may arise from domains in twinned orientation or from diffraction contrast due to local crystal bending. Bar, 15 nm.

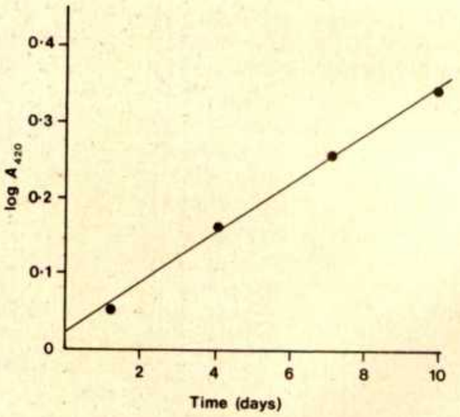


Fig. 3 Plot of log absorbance at 420 nm (A_{420}) against time for the intravesicular oxidation of Fe(II) solutions. The linear plot represents an overall first-order rate dependence. Samples were stored in the dark at 4 °C between readings. TEM studies showed minimal evidence of vesicle degradation over the course of the kinetic experiments.

These kinetic restrictions favour the formation of thermodynamically stable products such as geothite from Fe(III) solutions, rather than ferrihydrite which is formed under conditions of rapid precipitation. In addition, the retardation of the rate of intravesicular Fe(II) oxidation due to the slow increase in intravesicular *p*H favours the formation of mixed valence oxides such as magnetite whereas rapid oxidation at high *p*H in bulk solution generates hydrous iron(III) oxides such as goethite and lepidocrocite. The formation of intravesicular ferrihydrite rather than the bulk precipitated magnetite from 1:1 Fe(II)/Fe(III) solutions reflects a similar mediation of crystal growth by the membrane. The slow increase in intravesicular *p*H will allow oxidation and favour the preferential precipitation of the Fe(III) component at acid *p*H values.

Finally, particle sizes of the above intravesicular products are determined by the vesicle internal diameter and the encapsulation efficiency. In the experiments, assuming a simple salt solution and 100% encapsulation efficiency, ~2,000 Fe irons would be present per vesicle. This would result, for example, in intravesicular magnetite particles of 25 Å diameter. Our particles are slightly larger than this, which can be accounted for by additional ions arising from saturation binding of Fe at the headgroup sites (~1,500 on the inner surface, Fig. 2*a*). We also note that all the intravesicular particles are spherical or disk-shaped whereas bulk precipitates such as γ- and α-FeOOH are acicular. Nucleation of the intravesicular solids, induced at the lipid interface binding sites, is presumably such that the initial stages of crystal growth are partially constrained by the curvature of the vesicle membrane.

In conclusion, we have demonstrated the mediation of crystal-growth processes of iron oxides in the presence of organized membrane-bound assemblies. Future work will extend this aspect of membrane-mimetic chemistry to investigate the role of headgroup spacing, charge density and curvature on the nucleation and crystal growth of inorganic solids which are biologically and industrially relevant.

Received 1 September; accepted 17 October 1986.

1. Ford, G. C. *et al. Phil Trans. R. Soc.* **B304,** 551-565 (1984).
2. Mann, S., Perry, C. C., Webb, J. & Williams, R. J. P. *Proc. R. Soc.* **B227,** 179-190 (1986).
3. Mann, S., Frankel, R. B. & Blakemore, R. P. *Nature* **310,** 405-407 (1984).
4. Mann, S. *Struct. Bond.* **54,** 127-174 (1983).
5. Mann, S. *J. inorg. Biochem.* (in the press).
6. Birchall, J. D. *Trans. Br. Ceram. Soc.* **83,** 158-165 (1984).
7. Mann, S., Kime, M. J., Ratcliffe, R. G. B. & Williams, R. J. P. *JCS Dalton Trans.* 771-774 (1983).
8. Mann, S. & Williams, R. J. P. *JCS Dalton. Trans.* 311-316 (1983).
9. Stumm, W. & Morgan, J. *Aquatic Chemistry* (Wiley, New York, 1984).

A thylakoid processing protease is required for complete maturation of the lumen protein plastocyanin

Johan Hageman[*†], Colin Robinson[‡], Sjef Smeekens[*†] & Peter Weisbeek[*†]

[*] Department of Molecular Cell Biology and
[†] Institute of Molecular Biology, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands
[‡] Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

Plastocyanin, a photosynthetic electron carrier functional in the chloroplast lumen, is synthesized in the cytosol as a precursor (preplastocyanin) with an amino-terminal transit sequence[1]. This transit peptide contains the information specifying import into and routing within the chloroplasts and is removed in at least two steps[2]. An intermediate is observed in the stroma after the transport of preplastocyanin through the chloroplast envelope; mature plastocyanin is present in the lumen, after transport over the thylakoid membrane. We show here that the stromal processing protease[3] is not responsible for both processing events. It cleaves the precursor protein only to the intermediate size and a novel protease located in the thylakoids processes this intermediate to the mature protein. This second protease recognizes the processing intermediate but not the precursor. Thus plastocyanin import involves cleavage at the intermediate processing site mediated by the stromal protease and then cleavage at the mature processing site mediated by the thylakoid protease.

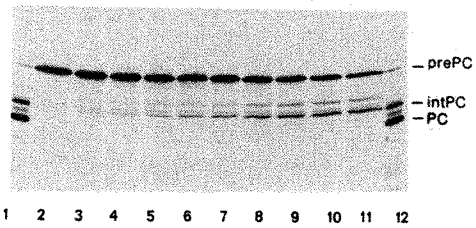


Fig. 1 Processing of plastocyanin precursor by stromal processing protease. Lane 1, import marker; lanes 2-11, incubation of plastocyanin precursor with partially purified protease for 0, 5, 10, 15, 20, 30, 40, 60, 90 and 120 minutes; lane 12, import marker.
Methods. Plastocyanin was synthesized in a wheat-germ system with RNA generated by *in vitro* transcription of the *Eco*RI-linearised pSPPC74 plasmid in an SP6 system[2]. The transcripts were capped by the addition of methylated cap ($m^7G(5')ppp(5')G$). The RNA was translated in a wheat-germ translation system in the presence of ^{35}S-methionine (1,000 Ci mmol^{-1})[14]. The stromal protease was concentrated ~350-fold from stromal extracts of pea chloroplasts as described[3]. The processing incubation at 27 °C contained 10 µl translation mix, 30 µl processing buffer (100 mM HEPES-KOH *p*H 8.5, 220 mM KCl, 5 mM $MgCl_2$) and 40 µl processing enzyme. At the times indicated, 8 µl aliquots were removed and put into a mixture of 32 µl H_2O and 40 µl sample buffer (125 mM Tris-HCl *p*H 7.6, 4% SDS, 5% sucrose, 5% 2-mercaptoethanol) and boiled for 2 min. The products were analysed by SDS-15% polyacrylamide gel electrophoresis using the Laemmli buffer system[15] and fluorography. Import marker: chloroplasts were isolated from young pea plants (cv Feltham First) as described[16] and finally taken up in import buffer (50 mM HEPES-KOH *p*H 8.0, 330 mM sorbitol) at a chlorophyll concentration of 1 mg ml^{-1}. In the import experiment 400 µl chloroplasts were incubated with 100 µl precursor plastocyanin translation mix, 25 µl 5× import buffer and 80 µl 200 mM methionine at 25 °C in the light[16]. After 60 min the chloroplasts were diluted with 10 ml import buffer and reisolated by centrifugation. The pellet was taken up in 600 µl sample buffer without protease treatment and boiled for 2 min. For each import marker lane, 15 µl of this mixture was used. The minor band between the processing intermediate and the mature plastocyanin may be an artefact as result of the length of the incubation.

Most chloroplast proteins are synthesized in the cytosol and subsequently transported into the organelle by a post-translational, ATP-dependent mechanism (for reviews see refs 4 and 5). The proteins are initially synthesized as larger precursors containing amino-terminal pre-sequences ('transit sequences') that are removed after import into the organelle. Transit sequences are believed to contain some, if not all, of the information specifying transport into the chloroplast and localization within the subchloroplast compartment[2,6,7]. Chloroplasts are composed of three membranes enclosing three discrete soluble phases—the intermembrane space, the stroma and the lumen. The proteins that are imported into the chloroplast have to cross one or more of these membranes; plastocyanin, a thylakoid lumen protein[8] must cross all three *en route* to its site of function. The import of preplastocyanin takes place in two stages[2], indicating that its transit sequence can be divided into an amino-terminal chloroplast-import domain and a thylakoid-transfer domain, each with its own processing site. We tried to determine whether the

```
        -66            -50      -40       -29     -20      -10     -1
prePC  MATVTSSAAVAIPSFAGLKASSTTRAATVKVAVATPRMSIKASLKDVGVVVAATAAAGILAGNAMA
PC58                                         MSIKASLKDVGVVVAATAAAGILAGNAMA
```

Fig. 2 Amino-acid sequence of the transit peptides of the wild-type plastocyanin precursor (prePC)[1] and the substitute intermediate (PC58). The amino acids are numbered relative to the mature processing site. The intermediate processing site is around the isoleucine residue at −27.

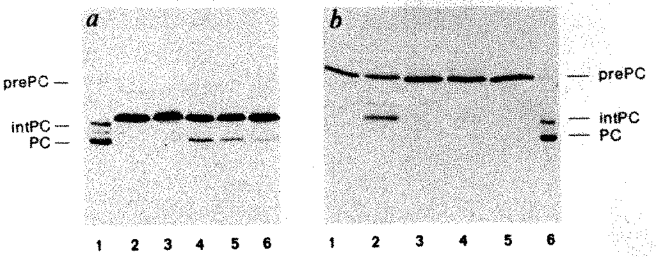


Fig. 3 A second protease is released from the thylakoids after Triton X-100 treatment. Incubation of PC58 (*a*) and plastocyanin precursor (*b*) with intact and Triton X-100-treated thylakoids. Panel *a*: lane 1, import marker; lane 2, PC58 translation mix; lane 3, incubation of PC58 with intact thylakoids; lanes 4–6, incubation of PC58 with thylakoids that were treated with 0.2, 0.5 and 1.0% Triton X-100 respectively. Panel *b*: lane 1, precursor plastocyanin translation mix; lane 2, incubation of preplastocyanin with intact thylakoids; lanes 3–5, incubation of preplastocyanin with thylakoids that were treated with 0.2, 0.5 and 1.0% Triton X-100 respectively; lane 6, import marker.

Methods. *In vitro* synthesis of PC58 was performed as described for plastocyanin precursor in Fig. 1 legend. The plasmid used is called pSPPC58. Thylakoids were isolated from isolated pea chloroplasts; intact chloroplasts at a chlorophyll concentration of 1 mg ml^{-1} in import buffer were pelleted by centrifugation and lysed by the addition of 20 mM Tris-HCl, *p*H 8.0, to the original volume. Aliquots of lysed chloroplasts were transferred to Eppendorf tubes and centrifuged in an Eppendorf centrifuge for 5 min at 4 °C. The pellet which contained the thylakoids was resuspended in the original volume of 20 mM Tris-HCl, *p*H 8.0. For the Triton X-100 treatment the thylakoids were pelleted again and taken up in solutions of 20 mM Tris-HCl, *p*H 8.0, containing the desired percentages of Triton X-100. The incubations (at 25 °C) contained 1 μl translation mix, 10 μl 20 mM Tris-HCl, *p*H 8.0, 10 μl processing buffer (see Fig. 1) and 70 μl untreated or Triton X-100-treated thylakoids. Reactions were stopped by adding 100 μl SDS-sample buffer followed by boiling for 2 min.

stromal processing protease is responsible for both the processing of the precursor to the intermediate and the intermediate to the mature form, or if not, what other proteases are involved.

Preplastocyanin was made by transcription *in vitro* of the *Silene* plastocyanin complementary DNA[2] in an SP6 vector[9], followed by translation of the RNA in a wheat-germ translation system. This procedure yields a single labelled band of precursor (Fig. 1 lane 2). The two minor lower relative molecular mass bands immunoprecipitate with plastocyanin antibodies (data not shown) and probably represent in-frame internal translation initiations at the methionine (ATG) codon at position −29 and at the isoleucine (ATC) codon at position −27 of the preplastocyanin transit sequence (see also ref. 2). The frequencies of these internal initiations vary between different wheat-germ systems (data not shown). The stromal processing enzyme was isolated from pea chloroplasts and purified ~350-fold[3]. As homologous size markers, mature plastocyanin and the processing intermediate were obtained from an *in vitro* import experiment with preplastocyanin (see Fig. 1 legend). The import marker lanes used in Figs 1, 2 and 4 show mature plastocyanin, a small amount of non-imported precursor and two intermediate bands the upper one of which represents the stromal processing intermediate[2].

Incubation of preplastocyanin synthesized *in vitro* with the partially purified stromal protease shows that the stromal processing protease is able to convert the precursor into a molecule the same size as the stromal intermediate (Fig. 1). It cannot be excluded that this processing occurs in two or more steps as the band between the precursor and the final processing product is consistently observed. The stromal protease however is not able to accomplish complete processing: even after prolonged incubation no mature protein is detected. (See also Fig. 4, lane 2.)

These results with *Silene* precursor plastocyanin differ from previous results obtained with barley and wheat plastocyanin[3]. We re-examined the processing of wheat plastocyanin precursor, using immunoprecipitated wheat plastocyanin mature protein as a reference during electrophoresis, instead of purified pea plastocyanin. This revealed that the protein that resulted from the processing with the stromal protease actually migrated more slowly than the immunoprecipitated mature plastocyanin from wheat (data not shown). The stromal protease therefore also processes the wheat plastocyanin precursor only to the intermediate size.

These results also imply that a second protease is responsible for the processing of the intermediate to the mature protein. Because the most likely location for this enzyme was the thylakoids, we incubated precursor protein with a thylakoid preparation. To release the putative protease from the thylakoids the membranes were disrupted by treatment with detergent[10]. Because of the possibility that precursor plastocyanin in addition to the precursor would not be recognized, a substitute intermediate protein was used to probe for the second protease. For the *in vitro* synthesis of this substitute intermediate a plastocyanin clone (PC58) was used which has had the initiator ATG codon removed by exonuclease digestion. Because of this, translation of *in vitro* transcribed *pc*58 RNA in a wheat-germ system starts at the first ATG, the methionine codon at position −29 downstream in the transit sequence (Fig. 2). Consequently PC58 protein is only a few amino acids longer than the stromal intermediate. PC58 is not imported into chloroplasts *in vitro* (data not shown). *In vitro* synthesis of PC58 yields one major labelled protein (Fig. 3*a* lane 2). The minor band above the PC58 protein probably represents an in-frame initiation at the isoleucine codon (position −55). The relative efficiency of initiation at this codon also depends on the wheat-germ extract used.

Formation of mature plastocyanin was not observed after incubation of PC58 protein or plastocyanin precursor protein with intact thylakoids (Figs 3*a*, lane 3; Fig. 3*b*, lane 2). PC58 was not processed at all, although the precursor was processed to the intermediate size. The latter processing was identical to that observed after incubation with stromal protease (Fig. 1) suggesting that the thylakoid fractions were not free from stromal contamination. Consistent with this, the SDS-polyacrylamide gels of the thylakoid preparation showed bands characteristic of ribulose bisphosphate carboxylase, the major stromal protein (not shown). PC58 is not processed by the stromal protease (data not shown) and therefore is not affected by stromal contamination.

Disruption of the thylakoid membranes by mild Triton X-100 treatment releases a protease activity that converts PC58 into a protein the same size as mature plastocyanin (Fig. 3*a* lanes 4–6). This processing activity is specific for the intermediate, as preplastocyanin is not processed at all under these conditions (Fig. 3*b* lanes 3–5), showing that the stromal protease contamination of the thylakoid fraction is completely inhibited by Triton X-100. The thylakoid protease is only moderately sensitive to this detergent. It can also be relased from the thylakoids by sonication (data not shown).

Finally, the combination of the stromal and the thylakoid protease is sufficient for the complete processing of preplastocyanin. After the incubation of the precursor first with the

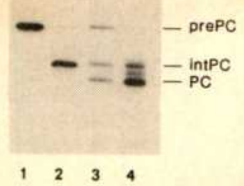


Fig. 4 Processing of plastocyanin precursor to the mature size by successive processing with a stromal protease and thylakoid protease. Lane 1, plastocyanin precursor translation mix; lane 2, incubation with stromal protease for 2 h; lane 3, incubation with stromal protease for 1 h followed by Triton X-100-treated (0.2%) thylakoids for 1 h; lane 4, import marker.

Methods. Plastocyanin precursor synthesized *in vitro* was incubated with partially purified stromal protease (see Fig. 1) for 1 h. Half was then incubated for 1 h with an equal volume of thylakoid vesicles treated with 0.2% Triton X-100 (see Fig. 3). The remainder was incubated for 1 h, after which samples were analysed by SDS-polyacrylamide gel electrophoresis followed by fluorography.

partially purified stromal protease and then with Triton X-100-treated thylakoids, a protein the same size as mature plastocyanin is detected (Fig. 4, lane 3). Prolonged incubation with stromal protease only resulted in a further decrease of the amount of precursor and an increase of the amount of intermediate (Fig. 4, lane 2), confirming that the stromal protease is not able to complete the processing. Processing of wheat preplastocyanin with this combination of proteases also results in complete processing to the mature size (data not shown).

The involvement of two distinct proteases in the processing of plastocyanin precursor to the mature size emphasizes the special nature of the transport route taken by nuclear-encoded proteins that are functional in the lumen. It also confirms the division of the plastocyanin transit sequence into a chloroplast-import domain and a thylakoid-transfer domain[2]. Partial processing of preplastocyanin by the stromal protease is consistent with the observation from the *in vitro* reconstitution experiments that the processing intermediate is found in the stroma, whereas the mature protein is found in the lumen[2]. The specificity of the thylakoid protease for the intermediate form suggests that this protease recognizes the plastocyanin transit sequence only after the removal of the chloroplast-import domain by the stromal protease. The determination of the exact location of the thylakoid protease is under investigation, and the first results on the purification show that the protease is water-soluble and retains activity on separation from the thylakoid membrane.

Plastocyanin, or more specifically the plastocyanin intermediate, is not the only protein that is imported from the stroma into the thylakoid lumen. A mechanism has to exist for lumen proteins made in the chloroplast too. By analogy with the proposed evolutionary origin of the chloroplast[11], this probably involves a prokaryote-like import mechanism. It will be interesting to determine whether plastocyanin and proteins synthesized in the chloroplast share import mechanisms or not, especially considering that the thylakoid-transfer domain remarkably resembles a prokaryote signal peptide[12]. The thylakoid protease would then function analogously to the signal peptidase[13].

This work is supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO) and by an EMBO short term fellowship to J.H.

Received 1 June; accepted 17 October 1986.

1. Smeekens, S., De Groot, M., Van Binsbergen, J. & Weisbeek, P. *Nature* **317**, 456–458 (1985).
2. Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K. & Weisbeek, P. *Cell* **46**, 365–375.
3. Robinson, C. R. & Ellis, R. J. *Eur. J. Biochem.* **142**, 337–342 (1984).
4. Ellis, R. J. *A. Rev. Pl. Physiol.* **32**, 111–137 (1981).
5. Cashmore, A. *Bio/Technology* **3**, 803–808 (1985).
6. Van den Broeck, G. *et al. Nature* **313**, 358–363 (1985).
7. Schreier, P. H., Seftor, E. A., Schell, J. & Bohnert, H. J. *EMBO J.* **4**, 25–32 (1985).
8. Boulter, D., Haslett, B. G., Peacock, D., Ramshaw, J. A. & Scawen, M. D. in *Plant Biochemistry II* (ed. Northcote, D.) 1–40 (University Press, Baltimore, 1977).
9. Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035–7056 (1984).
10. Siefermann-Harms, D., Ross, J. W., Kaneshiro, K. H. & Yamamoto, H. Y. *FEBS Lett.* **149**, 191–196 (1982).
11. Margulis, L. *Origin of Eukaryotic Cells* (Yale University Press, 1970).
12. Von Heijne, G. *EMBO J.* **3**, 2315–2318 (1984).
13. Benson, S. A., Hall, M. N. & Silhavy, T. J. *A. Rev. Biochem.* **54**, 101–134.
14. Anderson, C. W., Straus, J. W. & Dudock, B. S. *Meth. Enzym.* **101**, 635–644 (1983).
15. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
16. Grossman, A. R., Bartlett, S. G., Schmidt, G. W., Mullett, J. E. & Chua, N.-H. *J. biol. Chem.* **257**, 1558–1563 (1982).

Cholinergic neurones acquire adrenergic neurotransmitters when transplanted into an embryo

James Niles Coulombe & Marianne Bronner-Fraser

Developmental Biology Center and Department of
Developmental and Cell Biology, University of California, Irvine,
California 92717, USA

During development, cells become progressively restricted, until they reach their final phenotype. Differentiation was originally thought to be irreversible, but phenotypic plasticity has been observed in a variety of cell types, for example sympathetic neurones[1], the limb blastema[2] and some glial cell types[3]. A detailed description of the individual steps that lead to expression or reversal of phenotype is essential to understand the molecular events underlying cell differentiation. We examined whether ciliary neurones acquire adrenergic properties when exposed to a permissive embryonic environment. Cholinergic neurones were selectively labelled with a retrogradely transported marker and injected into chick embryos during active neural crest migration[4]. Four to five days after injection, some of the labelled neurones were found in 'adrenergic sites' and had developed catecholamine histofluorescence. The cells had thus accumulated adrenergic neurotransmitters even after differentiation into cholinergic neurones. This result shows that neurotransmitter plasticity occurs in cholinergic neurones and suggests that the neurotransmitter phenotype can be modified by the embryonic environment.

In the autonomic nervous system, both adrenergic and cholinergic neurones are derived from the neural crest[5]. In the differentiation of neural crest-derived neurones, the choice of neurotransmitter is a critical decision for proper autonomic function since the neurotransmitters used by sympathetic and parasympathetic neurones have antagonistic effects on their target organs. There is evidence that some neural crest cells may be fated to become cholinergic neurones; early migrating cranial neural crest cells have choline acetyltransferase activity[6] and acetylcholinesterase[7]. In addition, a monoclonal antibody to ciliary neurones recognizes a subpopulation of the neural crest cells in the mesencephalon[8]. After differentiation, the local environment may selectively promote survival of those neurones which contain the appropriate neurotransmitter, and/or instruct the developing cells to differentiate in the appropriate direction. Experiments showing that neurotransmitter choice in sympathetic neurones is labile even in post-mitotic cells[1] support the latter idea. Superior cervical ganglion cells (which normally produce catecholamines) produce acetylcholine[1,9–11] and form cholinergic synapses[12,13] when grown in the presence of a diffusible factor made by nonneuronal cells. Evidence for neurotransmitter plasticity in the autonomic nervous system is also found *in vivo* in the cholinergic sympathetic innervation to the sweat gland of the rat footpad[14–16] and in the cholinergic innervation of the iris following sympathectomy[17].

We examined neurotransmitter plasticity in cholinergic neurones, to ascertain whether cholinergic neurones are 'fated'

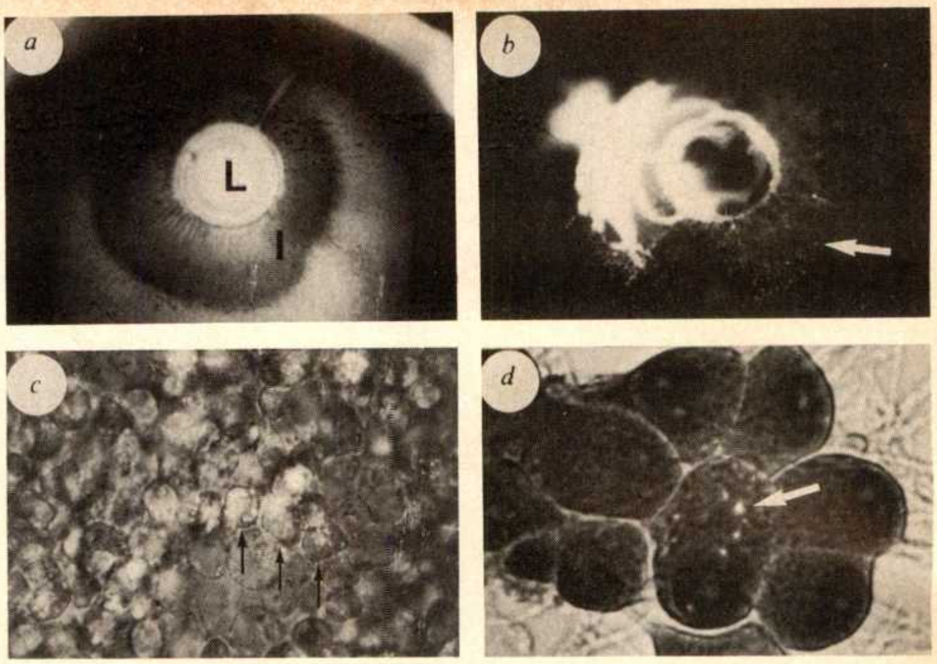


Fig. 1 Photomicrographs showing: *a*, *b*, the site of injection for retrogradely transported fluorescent latex microspheres in the iris region (I) of the eyeball. *c*, microsphere-labelled neurones in the ganglion. *d*, retrogradely labelled ciliary neurone. *a*, Transmitted light view of the iris region of a 6.5 day quail eye; L, lens (15×). *b*, Epifluorescence view of the same eye as *a*. Arrow, site where fluorescent latex microspheres filled the capillary bed of the ciliary body where axons from ciliary ganglion neurones terminate (15×). c, Combined fluorescence and bright field micrograph of a squashed ciliary ganglion showing fluorescent latex microspheres within neuronal cell bodies (596×). *d*, Bright field superimposed with fluorescence photomicrograph of a latex-based-labelled ciliary neurone (indicated by arrow) that has ChAT immunoreactivity (140×). All such neurones were found to be ChAT positive.

Methods. For labelling, the eyes were removed, leaving the ciliary ganglion in place at the back of the eyeball. The venous sinus (the canal of Schlemm), which supplies the ciliary ganglion and iris region with blood was severed ~3 mm proximal to the iris to avoid artefactual transport. Fluorescent latex microspheres (0.5–2.0 μm spheres, latex impregnated with rhodamine)[20] were pressure injected through the canal of Schlemm into the capillary bed that supplies the iris and ciliary body, the target sites for ciliary ganglion neurones. Because the venous sinus was severed, the neurones from the ciliary ganglion were only in contact with the latex microspheres at their growth cones and synapses, from which the microspheres were endocytosed and transported to the cell body. The eyeballs were incubated 8–10 h in a 95% oxygen, 5% CO_2 atmosphere at 37 °C in culture medium consisting of Eagle's minimum essential medium (Gibco) containing 15% horse serum, 10% chick embryo extract and 0.02% gentamycin sulphate. For ChAT immunocytochemistry, the microsphere-labelled ganglia were fixed in paraformaldehyde and picric acid, infiltrated with sucrose, and cryostat sectioned. Sections were dried onto gelatin coated microscope slides, and reacted overnight with primary antibodies against choline acetyltransferase (provided by Dr Miles Epstein) or tyrosine hydroxylase (Eugene Tech International Inc.). Antibody binding was made visible by treatment with biotin conjugated second antibodies, a biotin-avidin-horseradish peroxidase complex (Vectastain ABC), development with diamino-benzidine, and intensification with osmium. For cell counts, cell nuclei were stained with 4-6-diamidino-2-phenylindole dihydrochloride (DAPI).

to express only cholinergic neurotransmitters or labile with respect to neurotransmitter content. Previous investigators have transplanted cholinergic ganglia into presumptive adrenergic regions of the chick embryo[18]. In these transplants, the non-neuronal cells of the ganglia survived and synthesized catecholamines in the host adrenergic sites; the neurones, however, did not survive[19]. To test the transmitter plasticity of cholinergic neurones *in vivo* adequately, an experimental paradigm requires: (1) a transplantation method which supports neuronal survival; and (2) an unequivocal marker for the implanted cholinergic neurones. Here, we describe a new approach that meets these requirements.

To obtain a selectively labelled population of cholinergic neurones, we removed the eyes from 6.5 day quail embryos, leaving the ciliary ganglion in place at the back of the eyeball. Fluorescent latex microspheres were injected into a venous sinus supplying the iris and ciliary body, target sites for the neurones from the ciliary ganglion (Fig. 1). Microspheres were endocytosed by neuronal synapses and growth cones from ciliary neurones with processes at the target site, and were then retrogradely transported to the soma. The fluorescent latex microspheres, which were easily identified by their intense, punctate yellow-red fluorescence, have been shown to remain in the neuronal cell bodies for long periods of time and not to damage the neurone[20]. In 6.5 day old quail ciliary ganglia, ~15% of all cells were labelled with fluorescent microspheres. The labelled population represented ~20% of the neurones in the ganglia. The major advantage of this labelling technique is that only differentiated neurones, with processes at the target site, were labelled by retrograde transport.

To confirm that only neurones were labelled by the fluorescent microspheres, ciliary ganglia were dissociated and the cells were probed with an antiserum that recognized neurofilament proteins (Figure 2 and ref. 21). We classified a cell as labelled if it contained at least 5 fluorescent microspheres in its cytoplasm. Microsphere labelling was only found in cells containing neurofilament-like immunoreactivity. In addition, the microsphere-labelled cells had large cell bodies characteristic of neurones and many also had axonal processes.

An important assumption in our experiments is that the labelled neurones are cholinergic before transplantation. Four day old quail ciliary gangia have measurable choline acetyltransferase (ChAT) activity[22] and ciliary ganglion neurones *in vitro* form cholinergic synapses[23]. Recently, however, some tyrosine hydroxylase (TH) immunoreactive cells have been detected in avian ciliary ganglia[24]. Labelled ciliary ganglia were immunostained with antibodies against ChAT or TH to determine which of the neurotransmitter biosynthetic enzymes was present in labelled neurones. ChAT immunoreactivity (Fig. 1*d*) was found in all latex-bead-labelled neurones identified ($n > 1000$). In fact, all cells with the large cell bodies characteristic of neurones appeared to have ChAT immunoreactivity. Less than 1% of ciliary ganglion cells (estimated at 0.5% in chicken ciliary ganglia[25], and 0.8% in quail (unpublished data)) have TH immunoreactivity. No latex-bead-labelled neurones were detected among this population of TH-immunoreactive cells ($n = 104$).

Before implantation, the labelled ganglia were dissociated using conditions that promote neuronal survival in tissue culture[23]. The viability of the labelled neurones after dissociation was confirmed by their ability to regenerate severed axons and

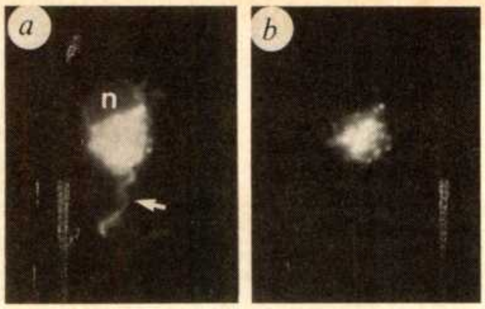


Fig. 2 Fluorescent photomicrographs of latex-bead-labelled ciliary ganglion neurones with anti-neurofilament antibodies. *a*, A ciliary ganglion neurone viewed with epifluorescence through a fluorescein filter has set neurofilament-like immunoreactivity. This cell has the large cell body and axonal process (arrow) characteristic of neurones. Due to fluorescence crossover, some latex microspheres could be seen through this filter set. Note the cell nucleus (n) 144×). b, The same neurone as in *a* viewed with a rhodamine filter shows numerous fluorescent latex microspheres in the cytoplasm (1440×). Diffuse signal is due to microspheres which were out of the plane of focus. Antibodies to neurofilament proteins were found to stain all cells containing fluorescent microspheres, suggesting that the microsphere-labelled cells were, in fact, neurones.

Methods. After labelling, ciliary ganglia were removed from the eyeball and cleaned of connective tissue and the ganglion capsule. The ganglia were then dissociated in 0.1% trypsin (Difco) in Modified Puck's Saline with glucose (MPG). Following trypsinization, the ganglia were mechanically dissociated by gentle trituration. For some experiments, the dissociated cells were filtered through a Nitex 15 μm screen. The dissociated cell suspension was fixed in 4% paraformaldehyde in 0.1 M phosphate buffer for 30 min at 22 °C. After washing, the cells were permeabilized in MPG containing 0.05% saponin with 5% goat serum for 30 min and reacted at 4 °C overnight with a rabbit primary antiserum (diluted 1:200) that specifically recognizes neurofilament proteins[21]. The cells were then washed and reacted with fluorescein-conjugated goat antibodies against rabbit immunoglobulin G (IgG) (diluted 1:20) for 1-2 hours at 37 °C.

to survive in culture. For some experiments, the ganglion cell population was enriched for neuronal cells by preplating[26]. The suspension containing labelled cholinergic neurones was implanted into the trunk somites of stage 12-18 chicken embryos using an injection technique[4], which reproducibly placed the cells on the ventral neural crest migratory pathway[27]. Embryos were incubated for an additional 4-5 days after injection, during which time some endogenous neural crest cells differentiated into catecholamine-containing cells[28]. The hosts were then rapidly frozen, processed for catecholamine-histofluorescence, and serially sectioned. During the 4-5 day incubation period, labelled neurones moved along the ventral neural crest pathway and collected around the dorsal aorta, where endogenous neural crest cells coalesce to form the adrenal medulla, sympathetic ganglia and aortic plexuses. Thus the labelled, dissociated ciliary ganglion neurones survived after implantation and were primarily located in or near sites normally populated by neural crest cells.

Injected embryos were examined for cells containing both fluorescent microspheres and catecholamines. We have identified, in nine separate embryos, a total of 33 microsphere-labelled ciliary ganglion neurones which also contained catecholamine histofluorescence (Figure 3*b*,*c*). Using a Feulgen stain, we have confirmed in several cases that the microsphere-labelled catecholamine-containing neurones also had the quail heterochromatin marker[29]. The number of labelled neurones containing adrenergic neurotransmitters represents ~5-10% of the microsphere-labelled neurones identified in section and ~50% of the labelled neurones that were incorporated into adrenergic structures. The other microsphere-labelled neurones were generally found near adrenergic cells but were not incorporated into adrenergic structures. Before injection no ciliary ganglion neurones contained catecholamines[25] or had a catecholamine uptake system[30]. Thus, some cholinergic neurones gained adrenergic neurotransmitters in the proper embryonic environment.

We observed a larger percentage of labelled neurones from ciliary ganglia in the adrenomedullary cords than in sympathetic ganglia. A similar distribution pattern to that of the labelled ciliary ganglion neurones is observed when inert latex beads are injected into the embryos[31]. Both the neurones and the beads may be moved to ventral sites by an environmentally controlled mechanism independent of cell motility. Only a fraction of the labelled neurones had catecholamine histofluorescence, and all of these were located in adrenergic structures such as the sympathetic ganglia or adrenomedullary cords. Localization in adrenergic structures may therefore be a prerequisite for the adrenergic phenotype. This suggests that the local environment, rather than the migratory process, influences the expressed phenotype of the neurones. A possible explanation for the small number of labelled neurones that were incorporated into host adrenergic structures may be that they are 'outcompeted' from these sites by endogenous neural crest cells. Because we have only examined the catecholamine content of the labelled ciliary ganglion neurones, our experiments cannot discriminate between uptake or synthesis of catecholamines. Neurones from ciliary ganglia *in vitro* do not have a catecholamine uptake system (refs 25, 30 and unpublished observations). Therefore, the labelled neurones have at least expressed biochemical systems for uptake and storage of catecholamines.

Control transplantations were performed to determine whether microsphere label could be passed to previously unlabelled cells after transplantation. Labelled ganglion cells were killed by repeated freeze/thawing, yielding microsphere-labelled ganglion-cell debris; the debris was then coinjected with an equal volume of a suspension of unlabelled live ciliary ganglion cells as in the experiments described above. In these controls, no cells containing multiple microspheres were identified. Individual microspheres were sometimes observed either in cells or extracellularly, but multiple microspheres were not acquired by unlabelled cells. This validates our selection of 5 microspheres as the criterion for labelling.

Thus, under appropriate conditions, cholinergic neurones can acquire an important characteristic of the adrenergic phenotype. Other investigators have observed that the environment encountered by neural crest cells can influence their subsequent differentiation. In quail/chick chimaeras, premigratory neural crest cells from presumptive cholinergic regions of the neural axis become adrenergic when grafted into the trunk[32]. In tissue culture, altering the environmental conditions can affect the differentiation of neural crest cells[33,34]. The culture environment can also induce explanted ciliary ganglion neurones to express some adrenergic characteristics; ciliary neurones *in vitro* can express immunocytochemically detectable tyrosine hydroxylase[24,25]. Similar results were observed *in situ* where some fibres from adult ciliary neurones apparently acquire tyrosine hydroxylase immunoreactivity after sympathectomy[17]. But, no catecholamines were detected in the ciliary neurones either *in vivo* after sympathectomy or *in vitro*, which contrasts with our findings.

Our results, together with the previous observations of the adrenergic to cholinergic transition[1,9-16], suggest that autonomic neurotransmitter plasticity is bidirectional. It remains to be determined whether the accumulation of catecholamines is due to uptake or synthesis, whether the transition can occur only in a defined period, whether cells reenter the cell cycle before expressing catecholamines and whether the labelled neurones can express both cholinergic and adrenergic traits simultaneously. Because transplanted cholinergic neurones can

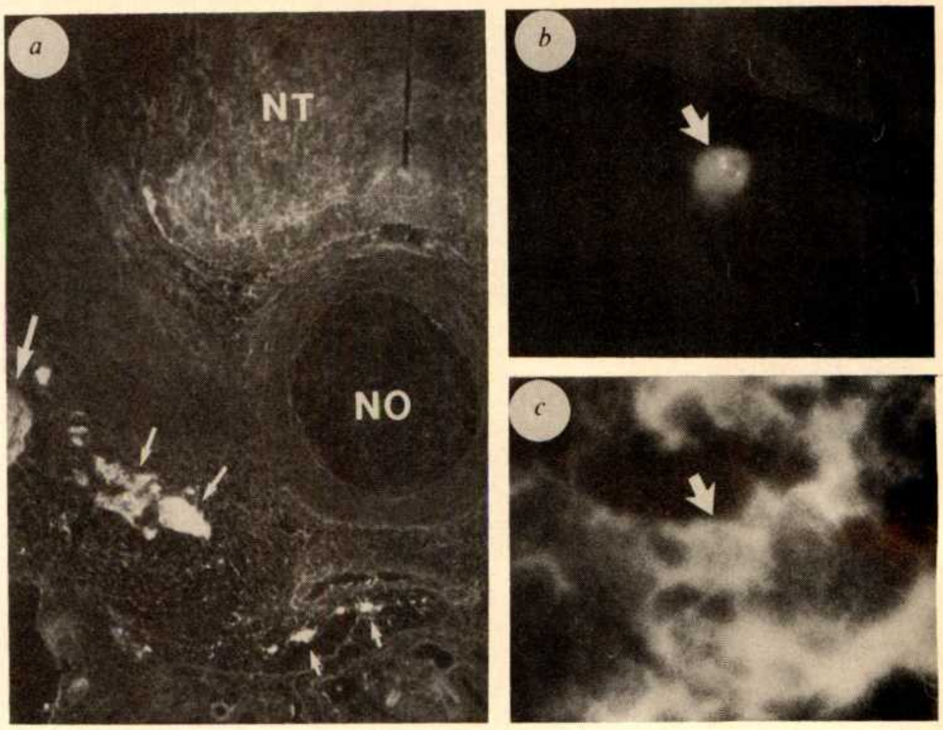


Fig. 3 Transverse sections through 7.5-day-old chicken embryos processed for catecholamine histofluorescence 5 days after labelled ciliary ganglion neurones were injected into the trunk somites. *a*, Low magnification of a transverse section through the trunk illustrating the sites where neural crest cells have formed catecholamine-containing derivatives. NT, neural tube; NO, notochord; arrows, sympathetic ganglia, adrenomedullary cells, and aortic plexuses (81×). *b*, A high power photomicrograph of the same embryo showing a retrogradely labelled neurone located in the adrenomedullary cords, viewed through a filter set designed for rhodamine fluorescence. Fluorescent latex microspheres, which have an intense punctate yellow-red fluorescence, can be visualized in the cell indicated by the arrow (890×). *c*, The same field as in *b* viewed through a catecholamine filter set, through which adrenergic cells are identified by their intense blue-green fluorescence. The microsphere-labelled cell (indicated by arrow) contains catecholamines and was part of a group of adrenergic cells located in the adrenomedullary cords (890×).

Methods. Embryos were rapidly frozen in liquid nitrogen-cooled isopentane, freeze-dried at −40 °C for ⩾4 days, equilibrated to room temperature and exposed to formaldehyde vapour for 1–2 h[35]. Fixed embryos were embedded under vacuum in paraplast, serially sectioned at 10 μm and placed dry onto albumin or acrylic adhesive coated microscope slides. Sections were covered with microscope immersion oil (Nikon) and examined at 800× under epifluorescent illumination. We confirmed that the same cell profile contained both microspheres and catecholamine histofluorescence using both rhodamine and catecholamine filter sets and examining the cell in a number of focal planes.

accumulate catecholamines, our results are consistent with a cell lineage model in which the neurotransmitter phenotype of autonomic neurones is strongly influenced by instructive cues from the local environment.

We thank Dr Scott Fraser for his helpful comments on the manuscript and Georgia Guillory and Jeffrey D. Pratt for technical assistance. This work was supported by USPHS Grant HD-15527-01 and by Basic Research Grant 1-896 from the March of Dimes.

Received 27 June; accepted 18 September 1986.

1. Patterson, P. H. & Chun, L. L. Y. *Proc. natn. Acad. Sci. U.S.A.* **71,** 3607–3610 (1974).
2. Brockes, J. P. *Science* **225,** 1280–1287 (1984).
3. Raff, M. C., Miller, R. H. & Noble, M. *Nature* **303,** 390–396 (1983).
4. Bronner, M. & Cohen, A. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1843–1847 (1979).
5. Le Douarin, N. M. *The Neural Crest* (Cambridge University Press, London, 1982).
6. Smith, J., Fauquet, M., Ziller, C. & Le Douarin, N. M. *Nature* **282,** 853–855 (1979).
7. Cochard, P. & Colty, P. *Devl Biol.* **98,** 221–238 (1983).
8. Barald, K. F. *Neuronal Development* (ed. Spitzer, N. C.) 110–119 (Plenum, New York and London, 1982).
9. Patterson, P. H. & Chun, L. L. Y. *Devl Biol.* **57,** 268–280 (1977).
10. Patterson, P. H. & Chun, L. L. Y. *Devl Biol.* **60,** 473–481 (1977).
11. Potter, D. D., Landis, S. C., Matsumoto, S. G. & Furschpan, E. J. *J. Neurosci.* **6,** 1080–1098 (1986).
12. Johnson, M., Ross, D. *et al. Nature* **262,** 308–310 (1976).
13. Furschpan, E., MacLeish, P., O'Lague, P. H. & Potter, D. *Proc. natn. Acad. Sci. U.S.A.* **73,** 4225–4229.
14. Landis, S. & Keefe, D. *Devl Biol.* **98,** 349–372 (1983).
15. Yodlowski, M., Fredieu, J. R. & Landis, S. C. *J. Neurosci* **4,** 1535–1548 (1984).
16. Leblanc, G. & Landis, S. C. *J. Neurosci.* **6,** 260–265 (1986).
17. Bjorklund, H., Hokfelt, M., Goldstein, L., Terenus, L. & Olson, L. *J. Neurosci.* **5,** 1633–1640 (1985).
18. Le Douarin, N. M., Teillet, M. A., Ziller, C. & Smith, J. *Proc. natn. Acad. Sci. U.S.A.* **75,** 2030–2034 (1978).
19. Dupin, E. *Devl Biol.* **105,** 288–299 (1984).
20. Katz, L. C., Burkhalter, A. & Dreyer, W. J. *Nature* **310,** 498–500 (1984).
21. Drager, U. C., Edwards, D. L. & Kleinschmidt, J. *Proc. natn. Acad. Sci. U.S.A.* **80,** 6408–6412 (1983).
22. Smith, J., Fauquet, M., Ziller, C. & Le Douarin, N. M. *Nature* **282,** 853–855 (1979).
23. Nishi, R. & Berg, D. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5171–5175 (1977).
24. Teitelman, G. *et al. J. Neurosci.* **5,** 29–39 (1985).
25. Iacovitti, I. *et al. Devl Biol.* **110,** 402–412 (1985).
26. Collins, F. & Dawson, A. *Proc. natn. Acad. Sci. U.S.A.* **80,** 2091–2094 (1983).
27. Bronner-Fraser, M. *Devl Biol.* **113,** 44–55 (1986).
28. Enemar, A., Falck, B. & Hakanson, R. *Devl Biol.* **11,** 268–283 (1965).
29. Le Douarin, N. M. *Expl. Cell Res.* **77,** 459–468 (1973).
30. Roher, H. & Sommer, I. *J. Neurosci.* **3,** 1683–1693 (1983).
31. Bronner-Fraser, M. E. *Devl Biol.* **91,** 50–63 (1982).
32. Le Douarin, N. M. & Teillet, M. A. *Devl Biol.* **41,** 162–184 (1974).
33. Howard, M. & Bronner-Fraser, M. *J. Neurosci.* **5,** 3302–3309 (1985).
34. Perris, R. & Löfberg, J. *Devl Biol.* **113,** 327–341 (1986).
35. Falck, B. & Owman, C. *Acta Univ. lung.* **2,** 4–25 (1965).

Variable and conserved neutralization antigens of human immunodeficiency virus

Robin A. Weiss*, Paul R. Clapham*,
Jonathan N. Weber*, Angus G. Dalgleish*,
Lawrence A. Lasky† & Phillip W. Berman†

* Institute of Cancer Research, Chester Beatty Laboratories,
Fulham Road, London SW3 6JB, UK
† Department of Molecular Biology, Genentech, Inc., 460 Point San
Bruno Boulevard, South San Francisco, California 94080, USA

Human immunodeficiency virus type 1 (HIV-1, HTLV-III/LAV), the retrovirus responsible for acquired immune deficiency syndrome (AIDS), shows a high degree of genetic polymorphism, particularly in the *env* gene[1–6]. We have examined sera from rabbits and guinea pigs immunized with gp130, a recombinant *env* glycoprotein[7], and sera from HIV-1-infected subjects[8], to test their capacity to neutralize a panel of genetically divergent HIV-1 isolates. The sera raised against recombinant antigen specifically neutralized the virus strain from which the *env* gene was cloned (HTLV-IIIB), but not an independent isolate (HTLV-IIIRF). One rabbit serum tested on seven isolates cross-neutralized two at lower titres. In contrast, human sera from Britain and Uganda, chosen for ability to neutralize HTLV-IIIRF, cross-neutralized six other

HIV-1 isolates. When serum and isolate were derived from the same subject, the serum was in some cases effective at slightly lower concentrations (higher titres). Human complement did not affect neutralization titres. These findings indicate that genetically diverse HIV-1 isolates carry both variable and widely conserved antigenic epitopes for neutralizing antibodies. The identification of shared epitopes may help the development of protective vaccines.

The production of a glycosylated recombinant envelope antigen of HIV-1 (gp130) by hamster CHO cells was recently described[7]. This recombinant antigen has at its amino-terminus 25 amino acids of the herpes simplex type 1 glycoprotein D, followed by amino acids 61–531 of the *env* gene of the HIV-1 isolate HTLV-IIIB (IIIB) representing most of the major envelope glycoprotein, gp120. This antigen is secreted by CHO cells. Rabbits and guinea-pigs immunized with partially purified gp130 in Freund's adjuvant neutralize the infectivity of the IIIB isolate and immunoprecipitate not only recombinant gp130 but also gp160 and gp120 of IIIB (ref. 7). In Table 1 we confirm the neutralization of the IIIB isolate by two different assays, but find no significant neutralization of a distinct, Haitian isolate, HTLV-IIIRF (RF).

To ascertain whether antisera raised against recombinant gp130 specifically neutralized only the HIV-1 strain from which the recombinant was derived, envelope pseudotypes of vesicular stomatitis virus (VSV) were prepared from seven different HIV-I isolates, and their sensitivity to neutralization was tested with rabbit anti-gp130 serum. In addition, twelve sera from British and Ugandan[9] subjects pre-selected for neutralizing activity to the RF isolate were screened, as was an enzyme-linked immunosorbent assay (ELISA)-positive Ugandan serum that was inactive for RF neutralization. Pre-immune rabbit serum, anti-HIV-1 ELISA-negative human sera and a VSV(HTLV-1) pseudotype were employed as negative controls.

Table 2 shows that the rabbit anti-gp130 serum specifically neutralized IIIB, cross-neutralized MN (ref. 1) and MA2 (ref. 10) isolates, but did not neutralize four other isolates. In contrast, the human anti-HIV-1 sera, whether British or Ugandan, cross-neutralized nearly all HIV-1 isolates, including two sera (British 23, 24) taken from subjects who were infected after receiving a single HIV-positive blood transfusion and who had no other

Table 1 Strain-specific neutralization of HIV-I by sera raised against recombinant *env* polypeptide gp130

Serum	Neutralizing titre HIV-1 IIIB	HIV-1 RF	VSV(HIV-1) IIIB	VSV(HIV-1) RF
Rabbit				
25	50	—	50	—
26	10	—	10	—
27	50	—	50	—
Guinea-pig				
1	10	—	10	—
2	50	—	50	—
3	50	—	50	—

Two methods were used to titrate the neutralizing activity of rabbit and guinea-pig sera to HIV-1 gp130, after heat inactivation at 56 °C for 30 min. The first assay measured surviving tissue-culture infectious doses (TCID) of HIV-1, when serum dilutions were mixed with 10^5 TCID in culture medium (RPMI-1640 with 10% fetal calf serum) for 1 h at 43 °C and dilutions of the reaction mixture were then added to 10^5 C8166 assay cells. The cultures were maintained for 7 days and the presence of large multinucleated syncytia was scored on day 4 and 7. HIV-1 end-point titrations ($TCID_{50}$) scored by appearance of syncytia in C8166 cells were equivalent to those scored by immunofluorescence of HIV-1 antigens in the cells or the presence reverse transcriptase in the medium of C8166 or H9 cells infected with HIV-1 (unpublished data). The neutralizing titre shown is the reciprocal of the highest dilution inactivating $\geq 90\%$ TCID of HIV-1. The second assay[8] used the neutralization of vesicular stomatitis virus (VSV) pseudotypes bearing the envelope glycoproteins of HIV-1. Serum dilutions were mixed with 200 plaque forming units (p.f.u.) of VSV(HIV-1) for 1 h at 43 °C before plating in 30 mm diameter wells on 2.5×10^5 adherent CCRF-CEM cells pretreated with DEAE-dextran. After adsorption for 1 h, 10^6 CCL64 mink cells were added, followed by an agar overlay 2 h later. VSV cytopathic plaques were counted 20–48 h after VSV(HIV-1) plating, and the neutralizing titre is expressed as the reciprocal of the highest serum dilution giving $\geq 80\%$ reduction of p.f.u. Pre-immune sera from the same animals had no neutralizing activity. The (−) sign indicates no significant neutralization with 1:9 serum dilution.

Table 2 Cross-neutralization of HIV-1 isolates by recombinant *env* anti-serum, and British and Ugandan human sera

Serum	ELISA	Neutralizing titres to VSV(HIV-1) RF*	IIIB†	MN†	ARV-2†	MA2†	RUT‡	Z129§	VSV (HTLV-1)
Rabbit									
pre-immune	—	—	—	—	—	—	—	—	—
anti-gp130	+	—	50	10	—	10	—	—	—
Human									
British 1	—	—	—	—	—	—	—	—	—
6	+	10	10	250	6,250	50	250	50	—
12	+	50	10	250	1,250	50	250	50	—
14	+	10	—	250	250	10	250	50	—
17	+	50	10	250	1,250	50	50	250	—
21	+	50	10	10	1,250	10	50	250	—
22	+	250	—	10	1,250	10	250	250	—
23	+	50	10	10	1,250	10	50	50	—
24	+	250	10	10	6,250	50	250	250	—
Uganda 2	—	—	—	—	—	—	—	—	—
17	+	50	10	50	1,250	50	10	250	—
29	+	—	—	—	—	—	—	—	—
40	+	250	10	—	1,250	10	250	250	—
42	+	50	10	10	6,250	250	250	50	—
48	+	50	10	10	1,250	250	250	50	—

VSV(HIV-1) pseudotypes were prepared as described[8] for all 7 HIV-1 isolates. They were propagated in H9 cells except for Z129 in the K12 clone of HUT78 cells. Serial dilutions of heat-inactivated sera were reacted with 200 p.f.u. of VSV(HIV-1) and the surviving p.f.u. titrated as described for Table 1. VSV(HTLV-1) was grown in HOS/PL cells as described[13], and was used as a control to demonstrate the specificity of VSV(HIV-1) neutralizing activity.

Countries of origin: * Haiti; † USA; ‡ Tanzania; § Zaire.

Table 3 Titration of serum neutralizing activity for autologous and heterologous HIV-1 isolates

		Neutralizing titre to VSV(HIV-1)						
Serum	ELISA	RF*	ARV-2†	ARV-4†	RUT‡	Z84§	Z129§	IIIB†
B1	—	—	—	—	—	—	—	—
RF	+	250	250	NT	250	50	NT	10
ARV-2	+	50	1,250	NT	50	NT	NT	10
ARV-4	+	50	1,250	250	250	10	NT	50
RUT	+	—	10	—	50	—	10	—
Z84	+	10	250	NT	50	—	NT	10
Z129	+	10	50	NT	50	NT	—	10

NT, not tested.
Countries of origin: * Haiti; † USA; ‡ Tanzania; § Zaire.

known exposure to HIV-1. The Ugandan serum that was negative for RF neutralization failed to neutralize any isolate. It appears that the sera of humans infected with HIV-1 are more cross-reactive for neutralization than the rabbit anti-serum raised against gp130. However, as the neutralizing titre of the rabbit anti-serum was only 1/50 for the homologous virus (IIIB), a low degree of cross-reactivity might not have been detected. Higher titre anti-gp130 sera would merit investigation.

Indirect immunofluorescence techniques were used on H9 cells productively infected with IIIB, RF and ARV-2 to test whether the rabbit anti-gp130 serum recognized non-neutralizing epitopes of HIV-1. Positive results were obtained with each isolate with anti-gp130. It is clear, therefore, that some common epitopes on the virion envelope identifiable by membrane immunofluorescence are not targets for neutralization.

Table 2 also shows that VSV pseudotypes prepared with certain HIV-1 isolates (such as ARV-2) are generally more sensitive to neutralizing sera than other isolates (such as IIIB). This property does not appear to correlate with absolute VSV(HIV-1) titre, as all pseudotype preparations contained $\sim 10^4$ plaque-forming units per ml.

We previously reported[8] that some sera had high titres for weak neutralization (50% plaque reduction) but that sera taken from the same patients as the isolates (LAV-1, CBL-1) did not show preferential neutralization of their own isolate. However, sequence analysis has indicated that the *env* sequences of the four isolates tested show <3% nucleotide divergence. We therefore tested the neutralizing activity of sera against HIV-1 isolates from the same patients, using HIV-1 isolates that are known to be genetically polymorphic (Table 3). The data indicate that some virus isolates are more sensitive to neutralization by their autologous serum (for example, RF) and that some weakly neutralizing sera (for example, RUT) neutralize the autologous isolate best. But Z84 and Z129 were not neutralized by autologous sera.

The strong membrane immunofluorescence seen with some sera which show weak or no neutralizing activity against HIV-1 (ref. 8) suggests that non-neutralizing antibodies bind to virion envelopes. We therefore investigated whether complement enhanced the neutralizing activity of antisera directed to HIV-1 envelope glycoproteins. Our results indicate that neutralizing titres are not significantly enhanced by human complement (not shown).

Our experiments reveal the presence of both common and strain-specific epitopes in the target antigens for antisera responsible for the neutralization of HIV-1. The studies with recombinant gp130 (ref. 7) and VSV(HIV-1) pseudotypes indicate that the major viral envelope glycoprotein, gp120, is a target for neutralizing antibodies, but that does not exclude gp41 and possible *gag* gene products. Where the *env* sequences are known (IIIB, ARV-2 and RF), they show up to 15% nucleotide sequence divergence[3,4], and it is likely that the Zairian and Tanzanian isolates are even more divergent[5]. HIV-1, like the lentiviruses of ungulates[11], shows considerable genomic polymorphism, particularly in those regions of the *env* gene encoding domains not characteristic of oncoviruses[6].

Although some of the genetic variability exhibited by HIV-1 is reflected in divergence of neutralization antigens, there are insufficient data as yet to determine to what extent the variants have been immunologically selected. The human sera may be more cross-reactive than those from animals immunized with non-infectious antigen because of exposure to different epitopes, (1) through processing of newly made viral antigens, (2) opportunities for infection by multiple strains, and (3) the possible generation of variants during the course of infection[12]. Two serum samples were from subjects who had been exposed to HIV-1 on only one occasion (blood transfusion), but that does not exclude the possibility of infection with more than one strain present in the donor's blood. It is most unlikely that all the British and Ugandan sera would cross-neutralize American, Haitian, European and African strains through exposure to multiple, variable epitopes of viral antigen. We therefore conclude that certain neutralization antigens of HIV-1 are conserved between the diverse strains.

The variation of HIV-1-neutralizing antigens is in striking contrast to HTLV-1, for which neutralizing sera and virus isolates from Japanese, British West Indian and United States subjects are indistinguishable, though distinct from HTLV-2 (refs 13, 14). However, the neutralizing serotypes of avian leukosis viruses show considerable variation even within a single envelope subgroup[15,16], and different isolates of the bovine leukosis virus share some neutralizing epitopes of the gp51 envelope glycoprotein[17].

Our demonstration of conserved neutralization antigens, possibly those close to the binding site of gp120 to the CD4 cellular receptor[18-20] or other functionally conserved domains, may be important in the development of vaccines that elicit humoral and cellular immunity to HIV-1 infection. The strain specificity of the sera raised against recombinant gp130 reported here and those raised against purified gp120 of HTLV-IIIB (ref. 21) may indicate immunodominant determinants specific to the IIIB isolate. Further studies using monoclonal antibodies, and sera raised against different *env* constructs are in progress. By using a panel of VSV pseudotypes enveloped by glycoproteins of genetically and immunologically variant isolates of HIV-1, antisera raised against recombinant glycoproteins or synthetic peptides can be rapidly screened to identify those eliciting antibodies cross-reactive for neutralization.

We thank R. C. Gallo and M. Popovic for providing HIV-1 isolates, HTLV-IIIB, RF, MN and serum RF, and HTLV-1 transformed C91/PL and C8166 cells; L. Thiry and J. A. Levy for ARV-2 and ARV-4 viruses and sera; and D. Zagury for Z84 and Z129 viruses and sera. Human sera from Uganda were provided by R. G. Downing and N. K. Sewankambo. D. Whitby provided technical assistance. This study was supported jointly by the Cancer Research Campaign and the MRC. J.N.W. is supported by a clinical research fellowship from the Wellcome Trust.

Received 25 July; accepted 13 October 1986.

1. Wong-Staal, F. *et al. Science* **229,** 759–762 (1985).
2. Benn, S. *et al. Science* **230,** 949–951 (1985).
3. Hahn, B. H. *et al. Proc. natn. Acad. Sci. U.S.A.* **82,** 4813–4817 (1985).
4. Starcich, B. R. *et al. Cell* **45,** 637–648 (1986).
5. Alizon, M., Wain-Hobson, S., Montagnier, L. & Sonigo, P. *Cell* **46,** 63–74 (1986).
6. Coffin, J. M. *Cell* **46,** 1–4 (1986).
7. Lasky, L. A. *et al. Science* **233,** 209–212 (1986).
8. Weiss, R. A. *et al. Nature* **316,** 69–72 (1985).
9. Serwadda, D. *et al. Lancet* **ii,** 849–852 (1985).
10. Webster, A. D. B. *et al. Lancet* **i,** 581–583 (1985).
11. Haase, A. *Nature* **322,** 130–136 (1986).
12. Hahn, B. H. *et al. Science* **232,** 1548–1553 (1986).
13. Clapham, P. R., Nagy, K. & Weiss, R. A. *Proc. natn. Acad. Sci. U.S.A.* **81,** 2886–2889 (1984).
14. Hoshino, H. *et al. Int. J. Cancer* **36,** 671–675 (1985).
15. Ishizaki, R. & Vogt, P. K. *Virology* **30,** 375–387 (1966).
16. Weiss, R. A. & Biggs, P. M. *J. natn. Cancer Inst.* **49,** 1713–1725 (1972).
17. Bruck, C., Portetelle, D., Mammerickx, M., Mathot, S. & Burny, A. *Leukemia Res.* **8,** 315–321 (1984).
18. Dalgleish, A. G. *et al. Nature* **312,** 763–767 (1984).
19. Klatzmann, D. *et al. Nature* **312,** 767–769 (1984).
20. McDougal, J. S. *et al. Science* **231,** 382–385 (1986).
21. Matthews, T. *et al. Proc. natn. Acad. Sci.* U.S.A. (in the press).

Cytotoxic T lymphocytes recognize influenza haemagglutinin that lacks a signal sequence

A. R. M. Townsend*, J. Bastin*, K. Gould† & G. G. Brownlee†

* Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK
† Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

A surprising feature of most cytotoxic T lymphocytes (CTL) responding to influenza infection is that they recognize the unglycosylated (non-transmembrane) proteins of the virus, including the nucleoprotein[1–4]. Recognition of cells that express nucleoprotein by CTL does not depend on a definite signal sequence within the protein[5], and the epitopes recognized can be defined with short synthetic peptides *in vitro*[6]. Haemagglutinin (HA), the major transmembrane protein of the virus, is recognized by a minor population of CTL from infected mice. We have deleted the sequence coding for the N-terminal signal peptide from a complementary DNA encoding HA of the H1 subtype. The signal-deleted HA is detected with antibodies as a short-lived, unglycosylated, intracellular protein. However, CTL raised to the complete molecule recognize cells expressing the signal-deleted HA and vice versa. These results cast doubts on the assumption that CTL recognize the HA molecule only after its insertion into the plasma membrane.

Intuitively, proteins that are naturally inserted into the plasma membrane should not require a special mechanism for recognition by CTL. The requirement for a hydrophobic signal peptide at the N-terminus of the HA molecule for its translocation into the lumen of the endoplasmic reticulum, insertion into the membrane, glycosylation and transport to the cell surface[7,8] formed the basis for our experiment. If CTL recognize haemagglutinin only after it has been inserted into the plasma membrane then cells expressing signal-deleted HA should not be recognized by HA-specific CTL. If an independent mechanism exists for presentation of HA in an appropriate form at the cell surface, then deletion of the signal sequence should not affect recognition by CTL.

The DNA encoding the HA signal sequence was deleted by oligonucleotide-mediated mutagenesis from a full-length DNA copy of RNA segment 4 from influenza A/PR/8/34 (H1 subtype, Mount Sinai strain, see ref. 2) as described[9] (Fig. 1*a*). We chose to express the signal-deleted HA from the 7.5K early promoter in recombinant vaccinia[10]. Such HA-vaccinia can prime mice efficiently for a haemagglutinin-specific CTL response[11,12], and express high levels of HA protein in cell lines *in vitro*. The recombinant viruses here are referred to as L$^-$H1-VAC (expressing the signal-deleted HA of the H1 subtype), H3-VAC (expressing the full-length HA of the H3 subtype from A/NT/60/68 virus) and H1-VAC (expressing the full length H1 haemagglutinin from the same promoter).

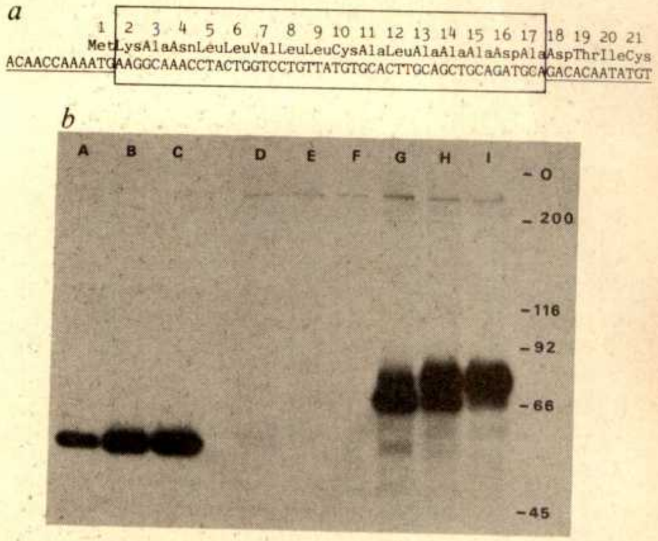

a
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
MetLysAlaAsnLeuLeuValLeuLeuCysAlaLeuAlaAlaAlaAspAlaAspThrIleCys
ACAACCAAAATGAAGGCAAACCTACTGGTCCTGTTATGTGCACTTGCAGCTGCAGATGCAGACACAATATGT
b
A B C D E F G H I
-0
-200
-116
-92
-66
-45

Fig. 1 *a*, Deletion of the H1 haemagglutinin (HA) signal sequence. *b*, Immunoprecipitation of signal deleted haemagglutinin. *a*, The N-terminal sequence of the H1 HA is from Winter *et al.*[13]. Sequence of the mutagenic oligonucleotide (24 bases) is underlined and the deleted segment coding for amino acids 2–17 is boxed. Mutagenesis was performed as described[9]. The mutant cDNA was sequenced by primer extension. The signal deleted H1 mutant and the full length H3 cDNAs were cloned into the vaccinia expression plasmid pGS62. Insertion into the vaccinia *tk* gene, isolation and purification of recombinant viruses followed standard procedures[10]. *b*, Murine L cells at 5×10^6 per ml were infected with 5–10 plaque forming units (PFU) of recombinant vaccinia virus per cell in RPMI containing 10% fetal calf serum (FCS) (RPMI/10) for 1.5 h at 37 °C. The cells were then washed twice in phosphate-buffered saline (PBS) and resuspended at 10^6 per ml in RPMI/5 for 3 h, washed twice in PBS and resuspended at 5×10^6 per ml in methionine-free RPMI, 5% FCS for 1 h before adding 120 μCi of ^{35}S-methionine for 30 min. They were then diluted to 10^6 cells per ml in RPMI/10 containing 4 mmol l^{-1} unlabelled L-methionine and the first aliquot of cells removed for lysis (time 0). The remainder were incubated at 37 °C and subsequent aliquots removed after 1 and 4 h. Immunoprecipitation was performed as described[5] on aliquots of 2×10^6 cells infected with L$^-$H1-VAC and NP-VAC (expressing the A/PR/8/34 nucleoprotein, used here as a specificity control), and 7×10^5 cells infected with H1-VAC. Reduced immunoprecipitates were electrophoresed on a 10% SDS polyacrylamide gel which was then treated with Amplify (Amersham), dried and exposed to pre-flashed X-ray film for 8 days. Lanes A–C, NP-VAC-infected cells immunoprecipitated with monoclonal antibody 469/6 (directed against nucleoprotein) at 0.1 and 4 h respectively after pulse labelling with ^{35}S-methionine. Lanes D–F, L$^-$H1-VAC-infected cells immunoprecipitated with H36-18-2 (anti-H1 haemagglutinin), time points as for A–C. Lanes G–I, H1-VAC-infected cells immunoprecipitated with H36-18-2, time points as for A–C.

The mutant gene, lacking the codons for amino acids 2–17 inclusive, should code for an unglycosylated protein of relative molecular mass 61.6K[13]. Figure 1*b* shows the results of a pulse-chase immunoprecipitation experiment using monoclonal antibody H36-18-2 that binds both the signal-deleted and full-length haemagglutinin. Immediately after a 30 minute pulse with ^{35}S-methionine, the mutant HA is weakly detectable in infected L cells at the expected position for unglycosylated HA (~60K, lane D). The full-length molecule migrates at about 68K, consistent with rapid core glycosylation (lane G). During the four hour chase period the mutant HA becomes undetectable, imply-

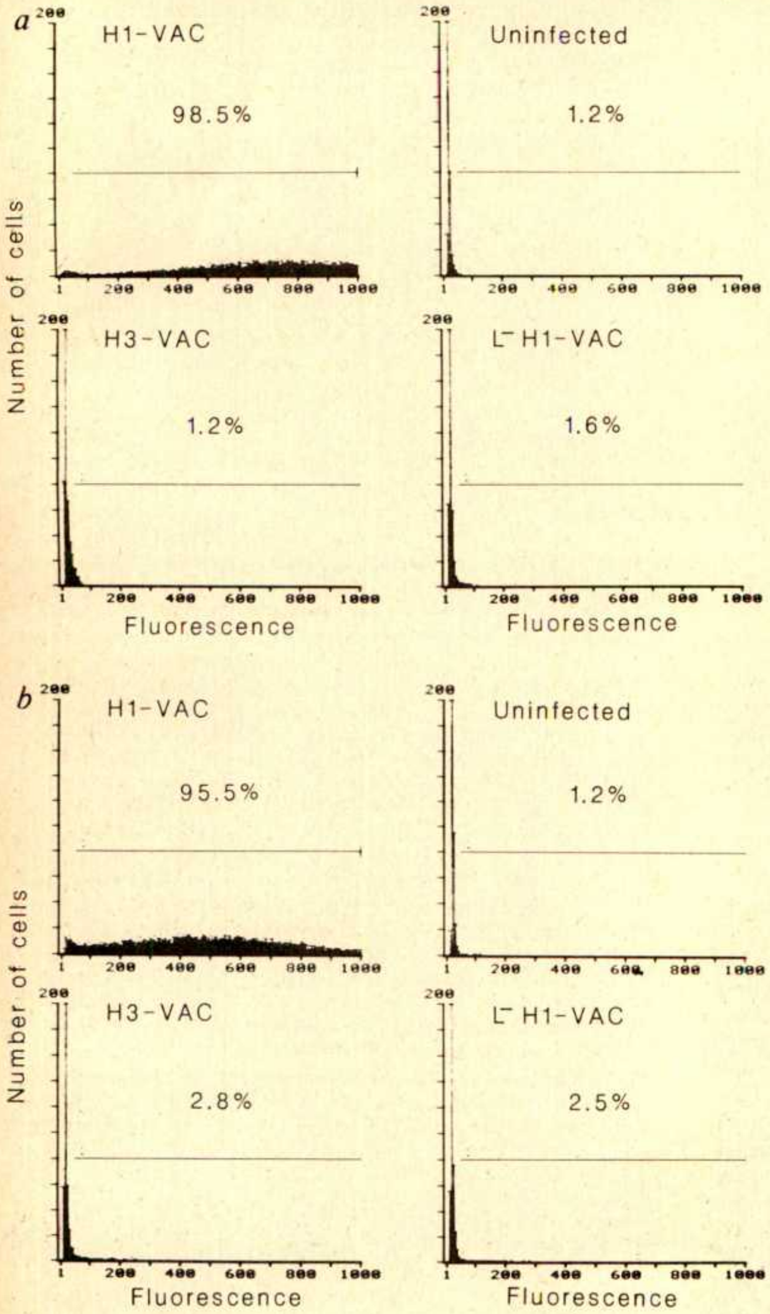


Fig. 2 Indirect immunofluorescence staining of L⁻H1-VAC- and H1-VAC-infected L cells. L cells were infected as described in Fig. 1 with either L⁻H1-VAC, H1-VAC, H3-VAC (included as a specificity control) or left uninfected. After 1.5 h cells were washed twice and resuspended at 10^6 per ml in RPMI/10 for 3 h at 37 °C. Aliquots of 5×10^5 cells were then stained by indirect immunofluorescence as described[5]. *a*, The first layer was H36-18-2 monoclonal antibody to H1 haemagglutin at 5 μg ml^{-1} after purification on protein A-Sepharose. The second layer was FITC-labelled affinity-purified goat anti-mouse antibody (Sigma) at 1:40 dilution. *b*, The first layer was hyperimmune rabbit anti-serum R401 raised to A/PR/8/34 virus diluted 1:2,000. The second layer was FITC-labelled affinity-purified goat anti-rabbit anti-serum (Sigma) diluted 1:100. The samples were analysed on an Ortho Cytofluorograf.

ing that it may be rapidly degraded (lanes D, E and F). In contrast the bulk of the full-length molecule progresses to a higher molecular mass of ~75K (lanes G, H and I). Control immunoprecipitations on cells infected with recombinant vaccinia expressing nucleoprotein (lanes A, B and C) gave a prominant band migrating at ~57K, as expected.

Figure 2 shows that the signal-deleted HA was not detectable at the cell surface with specific monoclonal (Fig. 2*a*) or polyclonal (Fig. 2*b*) antibodies. In contrast the full-length molecule was expressed at the surface of 98% of H1-VAC infected cells. In these experiments the target cells were infected for four and a half hours with recombinant vaccinia virus before being stained by indirect immunofluorescence. This is the equivalent of the mid time point in the ^{51}Cr release assay used to assess CTL recognition.

Recombinant vaccinia expressing the mutant and wild type HAs were then compared for their ability to prime CBA mice for a specific CTL response to the H1 subtype of influenza haemagglutinin. Animals primed *in vivo* with the L⁻H1-VAC were able to respond to wild-type HA expressed by A/PR/8/34 influenza virus *in vitro* as efficiently as mice primed with vaccinia expressing the full length HA. The response was shown to be specific because target cells expressing H1 but not H3 haemagglutinin were recognized (Fig. 3). Animals primed with H3-VAC (data not shown) did not make a detectable H1 haemagglutinin specific CTL response.

Finally, CTL from animals primed *in vivo* with A/PR/8/34 influenza virus were tested for their ability to recognize target cells expressing the signal-deleted HA. Most of the CTL in these polyclonal cultures are specific for conserved internal viral components that are shared by X31 (H3) and A/PR/8/34 (H1) influenza viruses. The subpopulation capable of recognizing haemagglutinin can be detected selectively with target cells expressing HA in the absence of other influenza proteins, after transfection or infection with a recombinant vaccinia[1,2,4,11,12].

Figure 4 (left panel) shows that the HA-specific CTL recognized target cells expressing the signal-deleted HA as efficiently as cells expressing the full-length molecule, despite the fact that

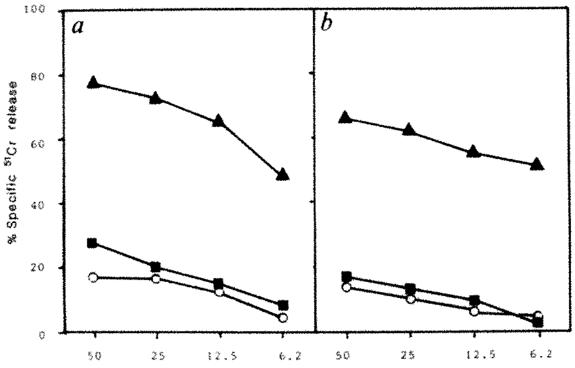


Fig. 3 The CTL response to infection by L⁻H1-VAC (*a*) and H1-VAC (*b*). Female CBA mice (obtained from NIMR Mill Hill), 3-6 months old, were primed intravenously by injecting 1×10^7 PFU of either L⁻H1-VAC or H1-VAC in 0.1 ml PBS into the tail vein as described[11]. After 12-14 days they were killed and their spleen cells stimulated *in vitro* with A/PR/8/34 influenza virus. After 5 days the responding cells were collected and tested by ^{51}Cr release assay on influenza-virus-infected L cells as described[1,2]. Responding CTL were incubated at the killer to target ratios indicated with ^{51}Cr-labelled target cells. Per cent specific ^{51}Cr release was calculated by (release by CTL − release in medium)/(2.5% Triton X-100 release − release in medium) × 100. Spontaneous release in medium from labelled L cells was 8-16% that released by Triton X-100. Target L cells were infected with; ▲, A/PR/8/34 (H1 subtype haemagglutinin); ■, A/NT/60/68 (H3 subtype haemagglutinin); ○, uninfected.

the appropriate antibodies could not detect HA on the surface of L⁻H1-VAC infected cells. Target cells infected with L⁻H1-VAC were not killed by a nucleoprotein specific T-cell clone, showing that recognition of the mutant HA was antigen specific. These experiments detect only CTL restricted through class I molecules of the major histocompatibility complex as the L cells used as targets do not express class II molecules.

These results are related to previous findings with nucleoprotein[5,6]. They show that CTL recognition of the haemagglutinin molecule (a typical glycosylated integral membrane protein[14]) does not depend on the presence of its signal sequence. Our findings suggest that an alternative route exists for the expression of HA at the surface of the cell in a form that is recognized by CTL. The lack of antibody binding to the L⁻H1-VAC-infected cell surface implies that the molecules recognized by CTL have disrupted tertiary structure. The pulse-chase experiment (Fig. 1) is consistent with this resulting from degradation of the mutant HA within four hours of its synthesis.

This conclusion is consistent with other recent published work. Wabuke-Bunoti *et al.* have shown that some HA-specific CTL can be induced *in vitro* with short synthetic peptides corresponding to regions of the H2 subtype haemagglutinin sequence[15]. Similarly, Yamada *et al.*[16] have described a recombinant protein produced in *Escherichia coli* consisting of the N-terminal 81 amino acids of the NS-1 protein fused to the C-terminal 210 amino acids of H1 haemagglutinin that was able to sensitize target cells for lysis by HA-specific CTL. We suggest, by analogy with the effect of denaturation on the recognition of proteins by T helper cells[17,18], that the tertiary structure of the HA fusion protein was disrupted enough to allow CTL recognition to occur.

Finally, Morrison *et al.*[19] have described Class I-restricted CTL that selectively recognized target cells in which new HA protein synthesis had occurred. In contrast their Class II-restricted CTL only recognized target cells that had been exposed to soluble HA in the surrounding medium. Our results argue that these findings can not be explained simply by the requirement for Class I-restricted CTL to recognize native HA inserted in the plasma membrane. We suggest that two pathways for presenting antigens to T cells exist, one dealing with proteins recently synthesized in the cytoplasm of an infected cell, the other concerned with endocytosed soluble proteins.

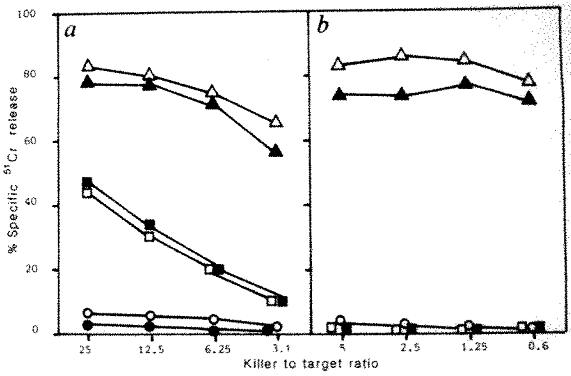


Fig. 4 CTL primed by full-length haemagglutinin recognize target cells expressing the signal-deleted molecule. *a*, Female CBA mice (3-6 months old) were primed intravenously with 100 haemagglutinin units of live A/PR/8/34 influenza virus. After 11 days spleen cells from primed animals were stimulated *in vitro* with A/PR/8/34 and collected for testing in the ^{51}Cr release assay as described in Fig. 3. Target cells used in this assay were L cells transfected with the Class I gene D^b (ref. 2). These cells were chosen so that the D^b-restricted nucleoprotein-specific CTL clone 11 could be used as an antigen specificity control. *b*, CTL clone 11 was isolated from a C57BL/10 mouse primed with influenza E61-13-H17 and restimulated with influenza X31 as described[20]. Clone 11 recognizes a conserved epitope on the nucleoprotein from human type A influenza viruses isolated since 1934 (J. Bastin, unpublished). Target cells were labelled with ^{51}Cr and infected with either recombinant vaccinia (5 PFU per cell) or influenza (200 μl of infectious allantoic fluid per 10^7 cells) for 1.5 h, washed twice, resuspended at 10^6 cells per ml in RPMI/10 for 3 h at 37 °C, washed twice and then used in the ^{51}Cr release assay as described[2]. Target cells were infected with; ■, H1-VAC; □, L⁻H1-VAC; ●, H3-VAC; ▲, A/PR/8/34 influenza (H1 haemagglutinin); △, X31 influenza (H3 haemagglutinin); ○, uninfected.

We thank G. Smith and B. Moss for the gifts of plasmid pGS62, recombinant vaccinia viruses H1-VAC and NP-VAC, and for initial help with vaccinia virus methodology. We are grateful to W. Gerhard, R. Webster and G. Schild for providing antibodies H36-18-2, 469/6 and R401 respectively; to N. Carter for analysing samples on the Ortho Cytofluorograf; to Claire Edmans for secretarial help and to the MRC for support.

Received 11 August; accepted 9 October 1986.

1. Zweerink, H. J., Askonas, B. A., Millican, D., Courtneidge, S. A. & Skehel, J. J. *Eur. J. Immun.* **7**, 630-635 (1977).
2. Townsend, A. R. M., McMichael, A. J., Carter, N. P., Huddleston, J. A. & Brownlee, G. G. *Cell* **39**, 13-25 (1984).
3. Yewdell, J. W., Bennink, J. R., Smith, G. L. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1785-1789 (1985).
4. Braciale, T. J., Braciale, V. L., Henkel, T. J., Sambrook, J. & Gething, M. J. *J. exp. Med.* **159**, 341-354 (1984).
5. Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42**, 457-467 (1985).
6. Townsend, A. R. M. *et al. Cell* **44**, 959-968 (1986).
7. Gething, M. & Sambrook, J. *Nature* **300**, 598-603 (1982).
8. Gething, M. & Sambrook, J. in *Genetics of Influenza Viruses* 169-191 (Springer, New York, 1983).
9. Zoller, M. J. & Smith, M. *Meth. Enzym.* **100**, 468-500 (1983).
10. Smith, G. L., Murphy, B. R. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7155-7159 (1983).
11. Bennink, J. R., Yewdell, J. W., Smith, G. L., Moller, C. & Moss, B. *Nature* **311**, 578-579 (1984).
12. Bennink, J. R., Yewdell, J. W., Smith, G. L. & Moss, B. *J. Virol.* **57**, 786-791 (1986).
13. Winter, G., Fields, S. & Brownlee, G. G. *Nature* **292**, 72-75 (1981).
14. Sabatini, D., Kreibich, G., Morimoto, T. & Adesnik, M. *J. Cell Biol.* **92**, 1-22 (1982).
15. Wabuke-Bunoti, M. A. N., Taku, A., Fan, D., Kent, S. & Webster, R. G. *J. Immun.* **133**, 2194-2201 (1984).
16. Yamada, A., Young, J. F. & Ennis, F. A. *J. exp. Med.* **162**, 1720-1725 (1985).
17. Streicher, H. Z., Berkower, I. J., Busch, M., Gurd, F. R. N. & Berzofski, J. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6831-6835 (1984).
18. Allen, P. M. & Unanue, E. R. *J. Immun.* **132**, 1077-1079 (1984).
19. Morrison, L. A., Lukacher, A. E., Braciale, V. L., Fan, D. & Braciale, T. J. *J. exp. Med.* **163**, 903-921 (1986).
20. Townsend, A. R. M. & Skehel, J. J. *J. exp. Med.* **160**, 552-563 (1984).

H-2-restricted cytolytic T cells specific for HLA can recognize a synthetic HLA peptide

Janet L. Maryanski*, Pietro Pala*, Giampietro Corradin†, Bertrand R. Jordan‡ & Jean-Charles Cerottini*

* Ludwig Institute for Cancer Research, Lausanne Branch, 1066 Epalinges, Switzerland
† Institute of Biochemistry, University of Lausanne, 1066 Epalinges, Switzerland
‡ Centre d'Immunologie INSERM-CNRS de Marseille-Luminy, Case 906, Marseille Cedex 9, France

It is generally accepted that T lymphocytes recognize antigens in the context of molecules encoded by genes in the major histocompatibility complex (MHC). MHC class II-restricted T cells usually recognize degraded or denatured rather than native forms of antigen on the surface of class II-bearing antigen presenting cells (reviewed in refs 1, 2). It has recently been shown[3] that short synthetic peptides corresponding to mapped antigenic sites of the influenza nucleoprotein (NP) can render uninfected target cells susceptible to lysis by NP-specific class I-restricted cytolytic T cells (CTL). These and earlier experiments[4] that showed specific recognition of NP deletion mutant transfectants suggest that class I-restricted recognition might also involve processed antigenic fragments. One important issue arising from these studies is whether the model applies not only to viral proteins that are expressed internally (such as NP) but also to antigens normally expressed as integral membrane proteins at the cell surface. We have recently isolated class I-restricted mouse CTL clones that recognize class I gene products of the human MHC (HLA) as antigens in mouse cell HLA-transfectants[5]. Here we show that these anti-HLA CTL can lyse HLA-negative syngeneic mouse cells in the presence of a synthetic HLA peptide. These results suggest that the model applies generally.

CTL clones were isolated as described elsewhere[5] from DBA/2 (H-2^d) mice immunized with syngeneic cells transfected with cloned HLA class I genes. A high efficiency transfection recipient[6] selected from mouse cell line P815 was used for DNA transfection. Anti-CW3 CTL clone 10.1 lysed P815 cells transfected with HLA-CW3, but not HLA-A3, -A24 or -B7 (ref. 5 and J.L.M., unpublished results). Recognition by this and other DBA/2 CTL clones directed against P815-HLA transfectants was shown to be restricted to H-2K^d (ref. 5). The antigenic determinant recognized by this CTL clone was mapped[7] to the second external (α_2) domain of HLA-CW3. Based on the amino-acid sequence of the α_2 domain of HLA-CW3 compared to the corresponding domain of HLA-A3, -A24 and -B7 and that of endogenous H-2^d class I molecules, we prepared a series of four synthetic peptides corresponding to 16-amino-acid segments of HLA-CW3 (Fig. 1).

```
    α1                                                α2                                               α3
     .....)(..........................................................................................)(.....
           90        100       110       120       130       140       150       160       170       180
            |         |         |         |         |         |         |         |         |         |
CW3   YYNQSEAGSHIIQRMYGCDVGPDGRLLRGYDQHAYDGKDYIALNEDLRSWTAANTAAQITQRKWEAAREAEQLRAYLEGLCVEWLRRYLKNGKETLQRAEHPKTH
A3              T  I       S   F    R D                    DM     K      H         D T         E        TDP
A24             TL M F     S   F    H Y          K         DM     K      HV   Q      T  DG     E         DP
B7              TL S               H  Y                    D                  R      E         E   DK E  DP
Kᵈ         KG   TF   F     S W      Q F    R         KT    D   L  R    Q GD  YY      E         EL N   L TDS  A
Dᵈ         AG   TL W A    ES        W F    C         KT    DM     R    Q GA  RD      E            NA  L TDP  A
Lᵈ         AG T TL W       S        E F    R         KT  F DM     R    Q GA  YY      E     H      NA  L TDS  A
```

Fig. 1 Comparison of the protein sequence of HLA-CW3[10], -A3[11], -A24[12], and -B7[13], and H-2K^d (ref. 14), -D^d (ref. 15), and L^d (ref. 16) from amino acid 83 to 188. Amino acid 181 of the HLA-CW3 sequence, erroneously given as G in ref. 10, has been corrected to R[17]. For H-2L^d, amino acids 121, 139 and 140 are given as C, S, and S, respectively in ref. 18. The four peptide sequences prepared synthetically are underlined.

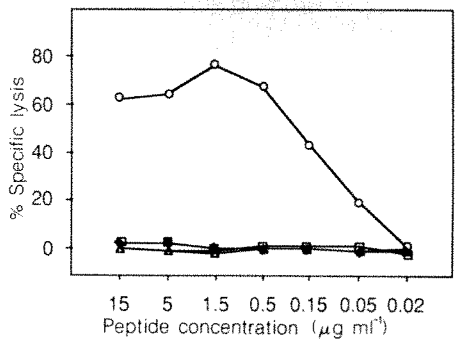


Fig. 2 Lysis of P815 target cells in the presence of peptide CW3(171-186) by anti-CW3 CTL clone 10.1. Synthetic peptides corresponding to HLA-CW3 amino acids 87-102 (△), 106-121 (◇), 151-166 (□) and 171-186 (○) were made according to the Merrifield method[19] as modified in ref. 20 using base-labile α-*N*-fluorenylmethoxycarbonyl protection. The peptides were cleaved from the resin with 90% aqueous trifluoroacetic acid, precipitated by addition of ether, redissolved in 10-20% acetic acid, chromatographed on a Sephadex G-25 column and eluted with 10-20% acetic acid. Peptides 87-102, 106-121 and 151-166 were then purified by reverse phase HPLC using a Permacoat C-6 column and elution with 0.1 M $NaClO_4$ in 0.1% H_3PO_4 buffer of *p*H 2 and a linear gradient of acetonitrile. Peptide 171-186 was purified by cation exchange HPLC using a Beckman Spherogel-TSK IEX 5.5 cm column and a linear gradient of NaCl in 0.01 M sodium phosphate buffer of *p*H 7.0. The peptides were lyophilized and dissolved in 0.7% sodium bicarbonate buffer. The HPLC-purified peptides gave the correct theoretical values of amino-acid composition and were >90% pure as judged by analytical HPLC. The purified peptides were diluted in Dulbecco's modified Eagle's medium (DMEM) containing 5% fetal bovine serum (FBS) and 100-μl volumes were added to wells of V-bottom microtitre plates. Cells (10^6) from P815-*tk*$^+$ clone 444/A1.1 (ref. 5) were labelled with 150 μCi sodium [^{51}Cr]chromate for 1 h at 37 °C and washed 3 times. Labelled P815 target cells (2×10^3 in 50-μl volumes) were added to wells containing peptides. 6×10^3 cells from anti-CW3 CTL clone 10.1 (ref. 5) in 50-μl volumes were then added to give an effector to target ratio of 3:1. After 4 h incubation at 37 °C the plates were centrifuged and 100 μl of supernatant was removed for counting. The % specific lysis was calculated as: (experimental − spontaneous release)/(total − spontaneous release) × 100. Spontaneous release in the absence of CTL or peptides was 14% of total counts. There was no effect on spontaneous release by any of the peptides. The specific lysis in the absence of peptides was 1%.

The CW3-specific CTL clone 10.1 lysed P815 cells incubated with peptide CW3 (171-186) very efficiently (Fig. 2). The minimum peptide concentration required for maximum lysis in this assay was 1.5 μg ml^{-1} (0.8 μM). The other three CW3 peptides tested were negative up to a concentration of 15 μg ml^{-1}. Another independently-derived H-2K^d-restricted DBA/2 anti-CW3 CTL clone (P1.1) apparently recognized the same HLA-

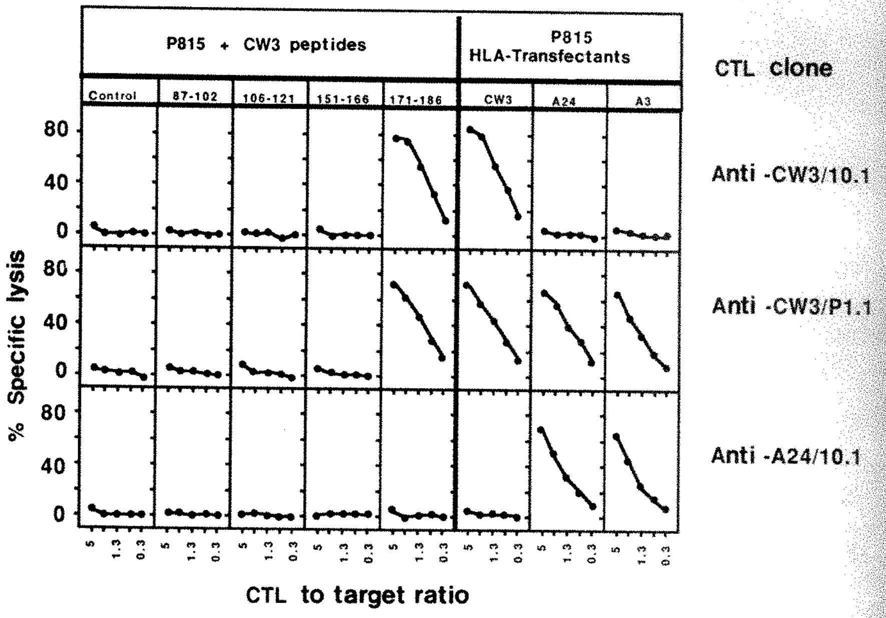


Fig. 3 Specificity of HLA-CW3 peptide recognition. Peptides CW3(171-186) and CW3(87-102) (100 μl at 10 μg ml^{-1} in medium containing 5% FBS) were added to conical microtitre wells. Cells (10^6) from P815 clones[5] transfected with the thymidine kinase (*tk*) gene only (clone 444/A1.1), with *tk* and HLA-CW3 genes (444/C9.3), with *tk* and HLA-A24 genes (452/D1), or with *tk* and HLA-A3 genes (504/2.1.4) were labelled with 150 μCi sodium [^{51}Cr]chromate for 1 h at 37 °C and washed 3 times. Labelled target cells (2×10^3 in 50-μl volumes) were added to wells that contained medium or peptides. Anti-CW3 CTL clones 10.1 and P1.1 were derived in independent experiments from DBA/2 mice immunized with P815 HLA-CW3 transfectant clone 444/C9.3, and anti-A24 CTL clone 10.1 was derived from DBA/2 mice immunized with P815 HLA-A24 transfectant cells[5]. Serial twofold dilutions of CTL in 50-4μl volumes were added to give the indicated CTL to target ratios and a final peptide concentration of 5 μg ml^{-1}. After 4 h at 37 °C the plates were centrifuged and 100 μl supernatant was removed for counting. Specific lysis was calculated as in Fig. 2.

CW3 synthetic peptide as clone 10.1 (Fig. 3). The fine specificity of the two CTL clones was not exactly identical, however. Clone P1.1 lysed P815-A24 and -A3 targets in addition to CW3 targets, unlike clone 10.1 that lysed only CW3 targets (Fig. 3 and ref. 5). Neither lysed HLA-negative P815-*tk*$^+$ cells. A third H-2K^d-restricted DBA/2 anti-HLA CTL clone that lysed HLA-A24 and -A3, but not -CW3 transfectants failed to react with any of the HLA-CW3 peptides, at concentrations of up to 30 μg ml^{-1} (Fig. 3 and J.L.M., unpublished results). Thus peptide recognition by these CTL clones (as measured by cytolytic activity) is specific, and antigenic sites for two different anti-CW3 CTL clones may be located within the same 16-amino-acid sequence.

These results clearly agree with those in ref. 3 showing specific cytolytic activity of anti-NP CTL in the presence of synthetic NP peptides. The reconstitution of specific target cell recognition with synthetic peptides strongly supports the hypothesis that class I-restricted T cells, like those restricted to class II antigens, recognize degraded or denatured antigenic fragments rather than intact antigen. It is noteworthy that HLA antigens, unlike the influenza NP, are integral cell membrane glycoproteins that can be easily detected on the cell surface with specific antibody[5,8]. Our results suggest, therefore, that recognition of antigenic fragments by class I-restricted T cells may be a general feature not limited to antigens that are virus-encoded and/or expressed primarily intracellularly.

Previous studies have shown that HLA class I gene products expressed in P815 mouse cells are involved in at least three categories of T-cell recognition. These include alloantigens recognized by alloreactive human CTL[8], HLA class I restriction elements recognized by human anti-viral CTL[9], and nominal antigens recognized by class I-restricted mouse CTL[5]. We have shown in this paper that the nominal HLA-CW3 antigen for two different class I-restricted CTL clones can be defined by a synthetic 16-amino-acid peptide. We have previously observed that lysis of P815-HLA transfectants by alloreactive human CTL but not by class I-restricted mouse anti-HLA CTL can be blocked by anti-HLA antibodies[5,8]. Together the results suggest that these human and mouse CTL recognize different forms of HLA that are presumably encoded by the same gene. It is possible that a membrane bound form of HLA in P815 HLA transfectants is recognized by alloreactive and by HLA-restricted antiviral human CTL, whereas a processed fragment of HLA is recognized as a nominal antigen by class I-restricted mouse CTL. The native and degraded forms of antigen may coexist on the cell surface. We believe P815-HLA transfectants will constitute an ideal system to investigate the nature of antigen complexes recognized by T cells.

We thank K. Mühlethaler and C. Vonnez for expert technical assistance and Dr H. R. MacDonald for helpful comments.

Received 15 September; accepted 10 October 1986.

1. Unanue, E. R. *A. Rev. Immun.* **2,** 395-428 (1984).
2. Grey, H. M. & Chestnut, R. *Immun. Today* **6,** 101-106 (1985).
3. Townsend, A. R. M. *et al. Cell* **44,** 959-968 (1986).
4. Townsend, A. R. M., Gotch, F. M. & Davey, J. *Cell* **42,** 457-467 (1985).
5. Maryanski, J. L., Accolla, R. S. & Jordan, B. *J. Immun.* **136,** 4340-4347 (1986).
6. Van Pel, A., De Plaen, E. & Boon, T. *Somatic Cell molec. Genet.* **11,** 467-475 (1985).
7. Pala, P., *et al.* in preparation.
8. Maryanski, J. L. *et al. Eur. J. Immun.* **15,** 1111-1117 (1985).
9. Gomard, E. *et al. Nature* **319,** 153-154 (1986).
10. Sodoyer, R. *et al. EMBO J.* **3,** 879-885 (1984).
11. Strachan, T., Sodoyer, R., Damotte, M. & Jordan, B. R. *EMBO J.* **3,** 885-894 (1984).
12. N'Guyen, C., Sodoyer, R., Trucy, J., Strachan, T. & Jordan, B. R. *Immunogenetics* **21,** 479-489 (1985).
13. Orr, H. T., Lopez de Castro, J. A., Parham, P., Ploegh, H. L. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 4395-4399 (1979).
14. Kvist, S., Roberts, L. & Dobberstein, B. *EMBO J.* **2,** 245-254 (1983).
15. Sher, B. T., Nairn, R., Coligan, J. E. & Hood, L. E. *Proc. natn. Acad. Sci. U.S.A.* **82,** 1175-1179 (1985).
16. Evans, G. A., Margulies, D. H., Camerini-Otero, R. D., Ozato, K. & Seidman, J. F. *Proc. natn. Acad. Sci. U.S.A.* **79,** 1994-1998 (1982).
17. Jordan, B. R., Lemmonier, F. A., Caillol, D. H. & Trucy, J. *Immunogenetics* **18,** 165-171 (1983).
18. Moore, K. W., Sher, B. T., Sun, Y. H., Eakle, K. A. & Hood, L. *Science* **215,** 679-682 (1982).
19. Marglin, A. & Merrifield, R. B. *A. Rev. Biochem.* **39,** 841-866 (1970).
20. Atherton, E., Logan, C. J. & Sheppard, R. C. *J. chem. Soc. Perkin* **I,** 538-546 (1981).

λ5, a new light-chain-related locus selectively expressed in pre-B lymphocytes

Nobuo Sakaguchi & Fritz Melchers

Basel Institute for Immunology, Grenzacherstrasse 487, Postfach,
CH-4005 Basel, Switzerland

The development from stem cells to pre-B cells, B lymphocytes and, finally, plasma cells and memory cells proceeds through various stages which have been defined by the genomic context in which immunoglobulin (Ig) heavy (H) and light (L) chain gene segments are found, as well as by their state of expression[1-7]. They have also been identified by surface marker analysis[5,8,9] and susceptibility to various stimuli regulating growth and differentiation[10]. We have searched for genes that are expressed at given stages in the B-lymphocyte development pathway and which might function to control this development at various stages. A com-

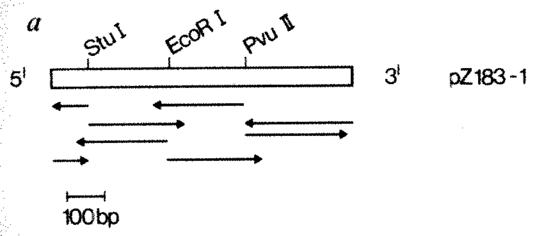


Fig. 1 Strategy for the sequencing and nucleotide sequence of the cDNA clone pZ183-1. The pZ183 cDNA clone was obtained as described elsewhere[11]. ^{32}P-labelled insert DNA was used to screen a pUC-13 cDNA library constructed from total poly(A)$^+$ RNA of 70Z/3 cells by the tailing method[25]. The restriction enzyme sites and the sequences are shown in *a*. The arrows indicate the direction and length of sequence determined by subcloning fragments into M13 phage vector and sequencing by the dideoxy-chain termination method[26]. The nucleotide sequence of the cDNA clone pZ183-1 is shown in *b*. The 5' to 3' orientation was deduced from the location of the poly(A) tail and the poly(A) attachment site (underlined). Nucleotide positions are numbered from the most 5' position of the insert cDNA fragment.

b

```
1        10        20        30        40        50        60        70        80        90       100
TTGGGTCTAGTGGATGGTGTCCACCACATACTTTCCCCAAGCTCACGCAGAAAGGACGAGAGCTGTGGGCCCTGGAGCTTCAGTGGGAACAACAGGCCTA

        110       120       130       140       150       160       170       180       190       200
GCTATGGCCCTTCCCGGCAGCTCCTGTTCCAGATCATCCCACGGGGAGCAGGTCCCAGGTGCTCGTTTTATAGGCTTCCATCTAAGCCCCAGTTTTGGTA

        210       220       230       240       250       260       270       280       290       300
TGTCTTTGGTGGTGGGACCCAGCTCACAATCCTAGGTCAGCCCAAGTCTGACCCCTTGGTCACTCTGTTCCTGCCTTCCTTAAAGAATCTTCAGCCAACA

        310       320       330       340       350       360       370       380       390       400
AGGCCACAGCTAGTGTGTTTGGTGAGCGAATTCTACCCTGGTACTTTGGTGGTGGACTGGAAGGTAGATGGGGTCCCTGTCACTCAGGGTGTAGAGACAA

        410       420       430       440       450       460       470       480       490       500
CCCAACCCTCCAAACAGACCAACAACAAATACATGGTCAGCAGCTACCTGACACTGATATCTGACCAGTGGATGCCTCACAGTAGATACAGCTGCCGGGT

        510       520       530       540       550       560       570       580       590       600
CACTCATGAAGGAAACACTGTGGAGAAGAGTGTGTCACCTGCTGAGTGTTCTTAGACCACAATCCTCCCTGAAGCCTCAGGGGCCTGGATCTGAAGTGCC

        610       620       630       640       650       660       670       680       690       700
AGAAAAAGTTGTTTTTTGTTTTGTTTTTTGTTTTTTTTCCCATTAACCATCTCACTGTCTTTCCTGTGCCTAATACTCAATAAATATCTTACCACCAACCA(24)
```

plementary DNA sequence called pZ183 was found in a library constructed from messenger RNA of the murine pre-B lymphoma cell line 70Z/3 which is selectively expressed in pre-B cells[11]. Here we report the nucleotide sequence of a cDNA clone (pZ183-1) containing 0.7 kilobases (kb) of the pZ183 gene. Part of this sequence shows strong homology to constant (*C*) and joining (*J*) region sequences of λ_1 L chains. Our findings define a new immunoglobulin L-chain-related locus, which we call λ5, that is selectively transcribed in pre-B lymphocytes.

During the somatic development of lymphocytes from precursor cells, gene segments in immunoglobulin gene loci are rearranged to form a complete variable (*V*) region exon which results in the translation into full immunoglobulin chains containing V- and C-region domains[12,13]. In the B-cell lineage rearrangement of heavy-chain diversity (D_H) segments to J_H segments precedes that of V_H to D_HJ_H segments[3,14,15]. V_κ is then rearranged to J_κ (refs 2, 16, 17) and, finally, V_λ to J_λ (ref. 18). The different steps of H, κ L and λ L rearrangments define different cellular stages of pre-B-cell development.

In our search for genes that are expressed at a given stage of pre-B-cell development, we had constructed a cDNA library from mRNA of (uninduced) 70Z/3 which was subtracted by a pool of mRNA molecules from a T-cell hybridoma[11]. We cloned the pZ183 cDNA which we found expressed selectively in pre-B cells at low to medium abundance[11]. It was present in 1 of 5,000 in the unsubtracted library and in 1 of 350 cDNA clones in a subtracted library from the same source of poly(A)$^+$ RNA of 70Z/3. Mature B cells, plasma cells, T cells, macrophages and a variety of cells from other lineages and tissues did not express detectable levels of the corresponding gene.

The cDNA clone with the longest (0.7 kb) insert, called pZ183-1, was obtained from a cDNA library constructed in the pUC-13 vector from a poly(A)$^+$ mRNA of uninduced 70Z/3 and sequenced as outlined in Fig. 1*a*. The nucleotide sequence is shown in Fig. 1*b*. Next to the poly(A) tail lies a poly(A) attachment site which orients the pZ183-1 sequence from 5' to 3'. The three longest open reading frames deduced from the cDNA sequence were considered as possible candidates for the polypeptide sequence: the longest frame between bases 1 and 553 (184 residues), the second between bases 9 and 281 (93 residues) and the third between bases 183 and 462 (93 residues).

We then searched a DNA sequence data bank[19] for possible evolutionary relationships of the pZ183-1 sequence to known DNA sequences. Strong homology was detected for positions

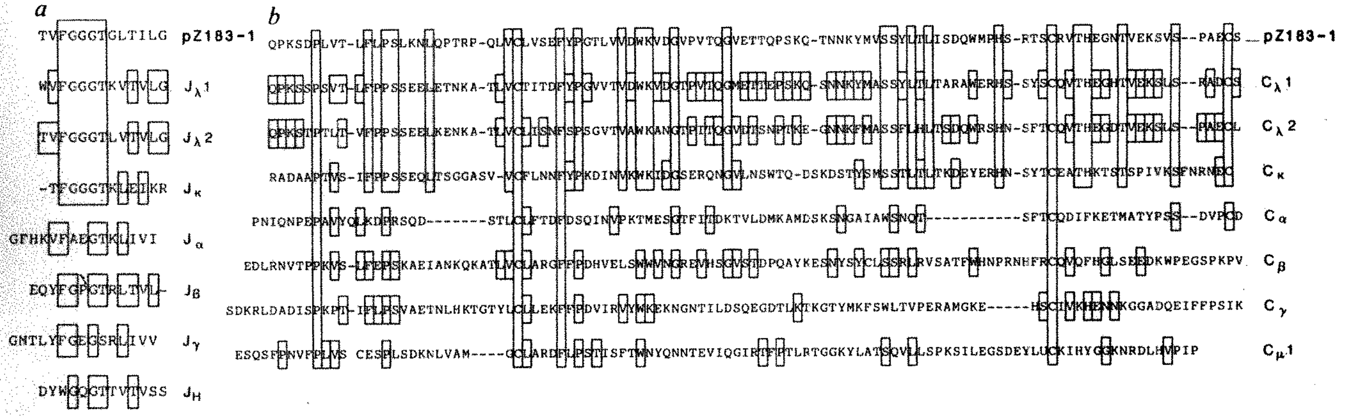


Fig. 2 Possible amino acid sequence of the *J*-like (*a*) and *C*-like (*b*) portions of pZ183-1 deduced from the nucleotide sequence in Fig. 1*b* in frame 1 and alignment with the corresponding sequences of λ1 L (refs 21, 27), λ2 L (refs 21, 27), κ L (ref. 21), J_H and $C_\mu 1$ (ref. 21) and J_α $C_\alpha 1$, J_β C_β and J_γ C_γ of the α, β and γ chains of the T-cell receptor[22,24,28]. See Table 1 for nucleotide positions corresponding to first and last deduced amino acid of the *J*-like and *C*-like sequences of pZ183-1. Those residues common between pZ183-1 and any other sequence are boxed. Residues conserved in all sequences or conserved in light chains are boxed together.

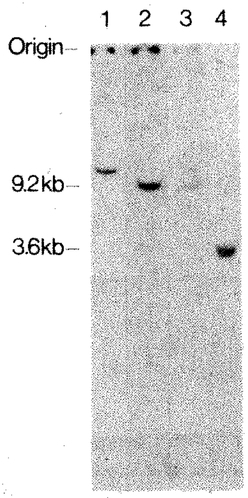


Fig. 3 Southern blot analyses of DNA from 70Z/3 cells. DNA was digested with the restriction enzyme *Hin*dIII, electrophoresed through a 1.0% agarose gel, blotted onto nitrocellulose filters and hybridized with the ^{32}P-labelled insert of the clone pZ183-1 (lane 1). Hybridization was carried out at 42 °C in 50% formamide and 5 × SSC. The filters were washed at 65 °C in 0.2 × SSC. After exposure to X-ray film, the same filter was washed in H_2O 0.1% SDS at 95 °C for 10 min and used repeatedly to probe successively with $C_\lambda 1$ [*Xho*I-*Pst*I fragment of pAB1-1 (ref. 27) (lane 2), both pZ183-1 and $C_\lambda 1$ (lane 3)] and $C_\lambda 2$ (450-bp *Sac* I fragment[23] (lane 4).

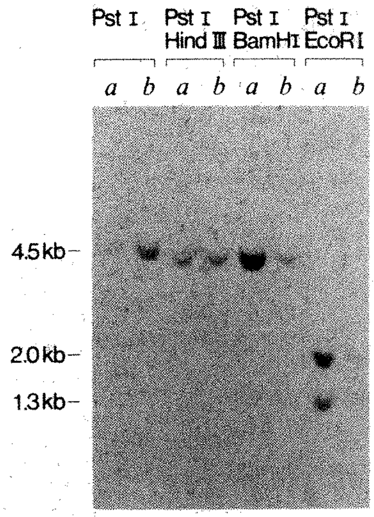


Fig. 4 Southern blot analyses of DNA from liver of (C57BL/6J × DBA/2)F_1 mice (*a*) and from 70Z/3 pre-B cells (*b*). DNA was digested with the indicated enzymes. The filter was hybridized with ^{32}P-labelled insert DNA of cDNA clone pZ183-1 under conditions described in Fig. 3 legend.

240–554 to the *C*-region nucleotide sequence of λ1 L chains. Equally strong homology was found for the adjacent 5′ positions 201–239 to the *J* segment sequence of λ1 (Table 1). The reading frame of the λ1 L-chain *J* + *C* nucleotide sequence was then imposed on the pZ183-1 sequence as a reading frame. This was found to be reading frame 1 (starting with nucleotide 1) of the sequence shown in Fig. 1*b*. Comparison with the λ1 nucleotide sequence shows that the first stop codon in frame 1 of the pZ183-1 sequence occurs at the same position as the first (functional) stop codon of λ1 L chains. Reading frame 1 is therefore the most likely candidate for a translatable sequence.

The amino acid sequence encoded by the *J*- and *C*-like regions of the pZ183-1 gene deduced from reading frame 1 is shown in Fig. 2 and compared with corresponding sequences of immunoglobulin and T-cell receptor proteins. Again, the strongest homology is detected with the *J* segment and *C* domain of λ1 L chains (Table 1).

Table 1 Homologies of pZ183-1 sequences with immunoglobulin light-chain genes and proteins

Sequence position of pZ183-1	L chain	Ref.	Homology with nucleotide sequence (%)*	Homology with deduced amino acid sequence (%)†
240–554	$C_\lambda 1$	21, 22	77.9	66.7
	$C_\lambda 2$	21, 22	70.2	61.1
	$C_\lambda 3$	22, 23	70.9	61.1
	$C_\lambda 4$	23	73.5	—
	C_κ	22, 24	22.5	30.8
201–239	$J_\lambda 1$	21, 22	67.5	66.0
	$J_\lambda 2$	21, 22	70.2	75.0
	$J_\lambda 3$	22, 23	67.6	50.0
	$J_\lambda 4$	23	70.2	—
	$J_1\kappa$	22,24	59.4	58.0

* Nucleotide sequence comparisons were done by a computer program and are shown as per cent matches per length.

† Alignment of the deduced amino acid sequence of pZ183-1 with mouse λ1, λ2, λ3 and λ4 L chains, as well as with κ L chains was done to maximize identities and structural homologies with pZ183-1, paying special attention to the alignment of certain key conserved residues.

Furthermore, pZ183-1 shows cysteine residues at the consensus positions characteristic of all immunoglobulin domains, and has the consensus sequence WVFGGGTKVTVLG characteristically encoded by *J* segments of the immunoglobulin genes. These results suggest that the 3′ portion of the pZ183-1 cDNA clone belongs to a λ L-chain-related gene locus which we call λ5 ($J_\lambda 5$, $C_\lambda 5$).

In contrast, the 5′ sequence adjacent to the $J_\lambda 1$-like sequence of pZ183-1 shows no strong homology to $V_\lambda 1$, any other known *V* segment or any other gene known to us. It does not contain a cysteine at the position where a consensus cysteine residue would be expected in V-region domains of immunoglobulins. The 5′ sequence adjacent to the *J*-like sequence of pZ183-1 also reveals no homologies to intron nucleotide sequences either 5′ of $J_\lambda 1$ or 3′ of $V_\lambda 1$. The segment cDNA clone is apparently too short to contain an initiation codon. We therefore cannot rule out the possibility that the pZ183-1 cDNA is derived from a sterile transcript.

The gene is carried in the genome of 70Z/3 cells in a *Hin*dIII fragment of >9.2 kb (Fig. 3, lane 1). Southern blot analysis shows that it is different in context from $C_\lambda 1$ (Fig. 3, lane 2) and $C_\lambda 2$ (lane 4). The unknown 5′ sequences and the $J_\lambda 1$- and $C_\lambda 1$-like 3′ sequences of pZ183-1 appear to be on the 4.5-kb *Pst*I fragment. The gene is on one *Pst*I/*Bam*HI and one *Pst*I/*Hin*dIII genomic fragment, each 4.2 kb long. Double digestion with *Pst*I and *Eco*RI yields two fragments of 1.3 and 2.0 kb (Fig. 4). No differences in the lengths of restriction enzyme fragments of DNA from liver and from 70Z/3 cells have been detected (Fig. 4). This indicates that the 5′ and 3′ regions of the pZ183-1 gene are near each other, that is, on one 4.3-kb fragment in germline DNA. We have therefore not found any sign that they are rearranged during pre-B-cell development. This also makes it unlikely that a chromosomal translocation or viral insertion occurred in close proximity to the pZ183-1 sequences in either the chemically induced 70Z/3 pre-B cell or the Abelson virus-induced pre-B-cell lines in which the same gene is expressed with the same size transcript of 1.2 kb[11].

The size of pZ183-1 transcripts, 1.2 kb, predicts a protein of similar size to L chains. With its penultimate cysteine at the carboxy-terminal end it could form homodimers (for example, L-chain dimers) or heterodimers. H chains could be partners in this heterodimer formation as soon as they are expressed in pre-B cells, and this may allow H chains to appear on the surface of pre-B cells, as previous experiments have indicated[9]. Such a surface deposition of the expressed H chains would significantly alter our present view of the development of the B-cell repertoire.

It had previously appeared that all the loci in mouse which encode antigen-binding immunoglobulin molecules of B lymphocytes had been identified, although considerable changes in the complexity of λ1 L-chain loci has been observed in wild mice[28]. However, our results suggest that immunoglobulin-related genes may exist which function only at an early stage in ontogeny and which so far have gone undetected.

We thank Ms Heidi Bächtold and Denise Richterich for technical assistance and Dr Klaus Karjalainen for critical reading of our manuscript. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche & Co., Basel, Switzerland.

Received 10 July; accepted 11 September 1986.

1. Tonegawa, S. *Nature* **302,** 575-581 (1983).
2. Maki, R., Kearney, J., Paige, C. & Tonegawa, S. *Science* **209,** 1366-1369 (1982).
3. Alt, F. W., Rosenberg, N., Lewis, S., Thomas, E. & Baltimore, D. *Cell* **27,** 381-390 (1981).
4. Perry, R. P., Kelley, D. E., Coleclough, C. & Kearney, J. F. *Proc. natn. Acad. Sci. U.S.A.* **78,** 247-251 (1981).
5. Coffman, R. L. & Weissman, I. L. *Molec. cell. Immun.* **1,** 33-38 (1983).
6. Paige, C. J., Kincade, P. W. & Ralph, P. *J. Immun.* **121,** 641-647 (1978).
7. Reth, M. G., Ammirati, P., Jackson, S. & Alt, F. W. *Nature* **317,** 353-355 (1985).
8. McKearn, J. P., Baum, C. & Davie, J. M. *J. Immun.* **132,** 332-339 (1984).
9. Kincade, P. W., Lee, G., Watanabe, T., Sun, L. & Scheid. M. P. *J. Immun.* **126,** 2262-2266 (1981).
10. Melchers, F., Andersson, J. & Phillips, R. A. *Cold Spring Harb. Symp. quant. Biol.* **41,** 147-158 (1977).
11. Sakaguchi, N., Berger, C. N. & Melchers, F. *EMBO J.* **5,** 2139-2147 (1986).
12. Hozumi, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73,** 3628-3632 (1976).
13. Sakano, H., Maki, R., Kurosawa, Y., Roder, W. & Tonegawa, S. *Nature* **286,** 676-683 (1980).
14. Sugiyama, H. *et al. Nature* **303,** 812-815 (1983).
15. Alt, F. W. *et al. EMBO J.* **3,** 1209-1219 (1984).
16. Lewis, S., Rosenberg, N., Alt, F. & Baltimore, D. *Cell* **30,** 807-816 (1983).
17. Burrows, O. D., Beck-Engeser, G. B. & Wabl, M. R. *Nature* **306,** 243-246 (1983).
18. Hieter, P. A., Korsmeyer, S. J., Waldmann, T. A. & Leder, P. *Nature* **290,** 368-372 (1981).
19. *GenBank, Genetic Sequence Data Bank* Release 35.0 (1985).
20. Scott, C. L., Mushinski, J. F., Huppi, C., Weigert, M. & Potter, M. *Nature* **300,** 757-760 (1982).
21. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. in *Sequences of Proteins of Immunological Interest* (NIH, Bethesda, 1983).
22. Hedrock, S. M., Nielsen, E. A., Kavaler, J., Cohen, D. I. & Davis, M. M. *Nature* **308,** 253-258 (1984).
23. Weiss, S., Meyer, J. & Wabl, M. R. *Eur. J. Immun.* **15,** 765-768 (1985).
24. Saito, H. *et al. Nature* **309,** 757-762 (1984).
25. Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning. A Laboratory Manual,* 230-242 (Cold Spring Harbor Laboratory, New York, 1982).
26. Sanger, F., Nicklen, S. & Coulson, A. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5463-5467 (1977).
27. Bothwell, A. L. M. *et al. Nature* **298,** 380-383 (1982).
28. Saito, H. *et al. Nature* **312,** 36-40 (1984).

Long-range restriction map around the Duchenne muscular dystrophy gene

Margit Burmeister & Hans Lehrach

EMBL, Postfach 102209, Meyerhofstrasse, 6900 Heidelberg, FRG.

Duchenne muscular dystrophy (DMD, reviewed in ref. 1) is an X-linked recessive disease affecting about 1 in 4,000 newborn boys. As in many other inherited diseases, the biochemical basis of the condition is unkown, and as yet there is no effective treatment. Translocations, deletions and other mutations leading to the DMD phenotype are distributed over a chromosomal area of large, but unknown size[2,3]. Using pulsed-field gradient gel electrophoresis[4,5], we have now determined restriction maps of a major fraction of this area, covering two regions of three million basepairs in total, and used it to determine the position of several probes linked to DMD. The maps establish physical distances between structural changes associated with the DMD phenotype and provide evidence for a CpG-rich island proximal to the area containing translocations and deletions associated with the DMD phenotype.

Using either molecular or genetic mapping, translocations, deletions or other mutations causing the DMD phenotype have been found distributed over an unusually large area of the X chromosome[2,3]. Although more complex genetic arrangements are possible, the simplest explanation is a single, but very large gene responsible for DMD. Two probes closely linked to the DMD gene and probably contained within have been cloned using different approaches. Subclones of a 200-kilobase (kb) cluster cloned around probe pERT87(DX164)[6] are deleted in about 6.5% of all DMD patients[3], suggesting that parts of the DMD gene might be deleted. Probe pXJ1.1 (ref. 7) is also assumed to be at the DMD gene, since it is derived from the chromosomal breakpoint of an X-chromosomal translocation found in a female DMD patient. As in other female DMD patients[2], the disease is assumed to arise due to disruption of the DMD gene by the translocation. Although extensive chromosomal walking has been carried out around both probes (P. Ray and R. Worton, and A. Monaco and L. Kunkel, personal communications), the two cloned regions could not be joined.

In contrast to most other genes, the size of the region implicated in the DMD phenotype is therefore too large to be easily accessible to standard molecular techniques. Recently, however, a new method, pulsed-field gradient (PFG) gel electrophoresis[4,5], has been developed, allowing the analysis of DNA fragments of 50-1,500 kilobases (kb); this range is not covered either by molecular cloning or by classical genetic approaches. We therefore decided that the analysis of the DMD locus would be considerably facilitated by mapping the positions of cloned probes relative to each other. Since these probes are associated with structural changes leading to a DMD phenotype, the positions of mutations leading to DMD are determined and estimations of a minimal gene size are possible.

We have analysed five probes closely linked to and flanking the DMD gene using the PFG technique. In addition to probes pERT87 and pXJ1.1, which are very close to or at the DMD gene, the closest available polymorphic markers on the distal side, B24 (DXS67) and C7(DXS28)[8], about 5-15 cM from DMD[9], and L1-4 (DXS68)[10], a non-polymorphic marker localized between pERT87 and C7 (ref. 9), were included. Deletion analysis[9,11-13] indicates the probe order pXJ1.1/pERT87/L1-4/C7/B24, and maps the genes for glycerol kinase and adrenal hypoplasia between pERT87 and C7.

To establish a physical map in this region, DNA was cleaved with different restriction enzymes cutting rarely in the mammalian genome, separated on PFG gels, and transferred to nylon filters for analysis. Filters were hybridized successively with each of the five probes. Probes pERT87-30 (the most distal subclone of the pERT87-cluster) and pXJ1.1 detected common restriction fragments; B24 and C7 were similarly shown to be physically linked. To generate a restriction map, double digests with combinations of enzymes were performed and analysed (see Fig. 1). The results of this analysis are summarized in Table 1. In a manner analogous to conventional restriction mapping, the data were used to deduce restriction maps covering hundreds of kilobase pairs. The restriction maps constructed (see Fig. 2 and legend for details) show that probes pXJ1.1 and pERT87-30 are at least 100, but no more than 400 kb apart. This result also gives the distance between the X/21 DMD translocation[7] (of which pXJ1.1 is a junction fragment) and common DMD deletions (100 kb proximal of pERT87-30, around pERT87-8)[3] as maximally 300 kb (see Fig. 2). The size of the DMD gene is probably larger, however, since some mutations map distal to pERT87-8.

Overlapping clones which extend ∼200 kb proximal to pERT87-30 have been isolated without joining the segment cloned around pXJ1.1 (ref. 3 and A. Monaco and L. Kunkel, personal communication); thus the maximum additional walking distance between the two cloned segments is now no more than 200 kb, and more likely to be much less.

B24 and C7 are the closest published markers distally flanking the DMD gene[9]. They are also of particular interest, since deletions which produce a syndrome of adrenal hypoplasia, glycerol kinase deficiency and DMD, map proximal to these markers[11-13]. Our results show that these two probes lie within a distance of 250–700 kb from each other. We could not show physical linkage between these probes and the pERT87/pXJ1.1 cluster (Fig. 2, Table 1). In an attempt to detect physical linkage between the two clusters, probe L1-4, which had been localised between C7 and pERT87-30 by deletion analysis[9,11], was hybridized to the same filters. No common restriction fragments were

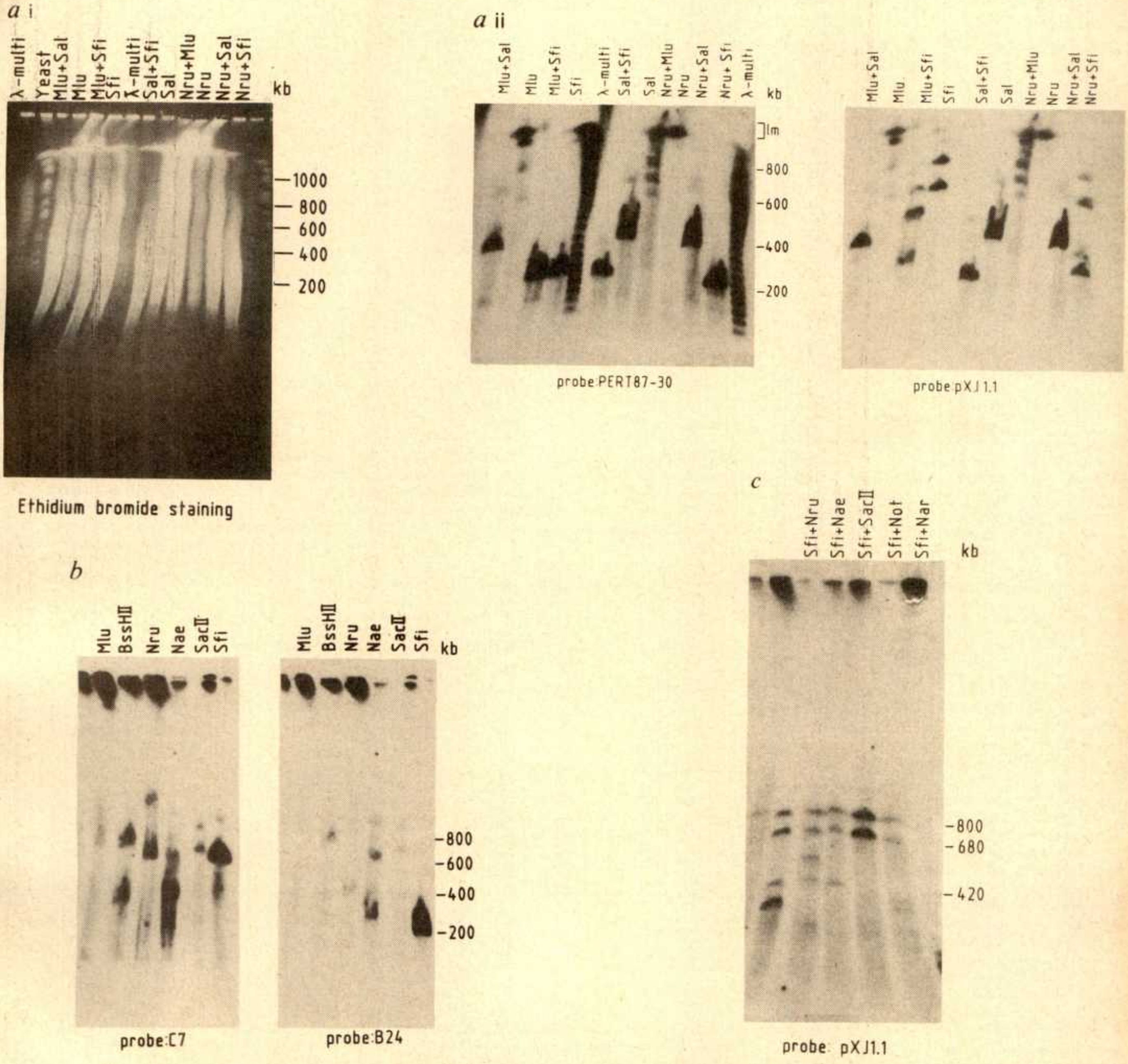


Fig. 1 Hybridizations to PFG gels. *a*, Linking probes pXJ1.1 and pERT87-30; *b*, common bands detected by probes B24 and C7; *c*, indications for an CpG-rich island near probe pXJ1.1. *a*, Ethidium bromide staining and autoradiograms of a gel run for 38 h at 15 °C with alternating switching of 48 to 90 seconds; These conditions give good resolution in the 200–600 kb range, but no resolution above about 900 kb. Size markers were both yeast chromosomes[4] and λ EMBL3 (ref. 18) polymers (λ-multi), which were generated by annealing the *cos* ends in agarose blocks. The unit size of EMBL3 is 43.2 kb. The autoradiograph at left shows the hybridization of λ DNA in addition to the probe used to visualize the size markers. The region of limiting mobility, lm, is the position of DNA fragments bigger than about 900 kb. *b*, This gel was run for 39 h at 9 °C with alternating switching times of 30 to 130 s, resolving also DNA fragments above 1,100 kb. Enzyme digests and probes are as indicated in the figure. Yeast chromosomes run on the same gel were used as size markers. c, Double digests of DNA with *Sfi*I and other enzymes as indicated were run on a PFG gel for 40 h at about 20 °C with alternating pulses of 30 and 130 s. *Sfi*I + *Not*I is equivalent to *Sfi*I alone. A common 390-kb fragment is detected with four enzymes in double digests with *Sfi*I. (*Sfi*I/*Bss*HII digests result in a similar fragment not shown in this figure). A 390-kb fragment is also observed when DNA is digested with *Nae*I alone. Therefore, the 390-kb fragment seen in the *Sfi*I + *Nae*I lane does not argue for the localization of an NaeI site close to the *Bss*HII/*Sac*II/*Nar*I/*Nru*I cluster.

Methods. Preparation of DNA for PFG was essentially as described[19]. Exact conditions for DNA preparation, running conditions and blotting will be published elsewhere[20]. A detailed protocol is available on request. The cell line LCL127 (48, XXXX, a gift from U. Francke, Yale) was used as a source for DNA in all cases. The source of the DNA is important because of observed methylation differences. Many sites seen in this cell line were not detectably cut in male blood cell DNA. Restriction enzymes were from New England Biolabs and were used as recommended by the manufacturer. Pulsed field gel electrophoresis[4] was carried out in a slightly modified box. The field was inhomogenous in both directions, the voltage gradient $10\ \mathrm{V\,cm^{-1}}$. Pulse times are as indicated above. After staining and photography, gels were alkaline-blotted onto Gene Screen nylon membranes. Inserts were isolated from low melting point agarose gels labelled by random oligo-priming[21] to a specific activity of around $1\text{-}4\times10^{9}$ d.p.m. μg^{-1} DNA. Hybridisations were carried out as described[22] for at least 24 h. Exposure was for 1 to 7 days at −70 °C with Kodac-X-Omat films preflashed once. Before reprobing, filters were stripped with 2 mM Tris/HCl *p*H 8, 0.1% sodium dodecyl sulfate at 80 °C. When the strongest band from the previous hybridization was not visible after reprobing, stripping was considered complete.

Table 1 Summary of hybridization results

Enzymes	B24	C7	Probes L1-4	pERT87-30	pXJ1.1
*Sfi*I	$\underline{200}$	$\underline{650}$ (1000)	$\underline{220}$ 350	$\underline{230}$	$\underline{680}$ 800
*Mlu*I	—	—	—	630 850 >950	630 850 >950
*Bss*HII	(350) 750	400 750	(850)	—	—
*Sac*II	(700) (900)	(700) (900)	(800)	(710)	(710)
*Sal*I	ND	ND	$\underline{300}$	$\underline{400}$ (650)	$\underline{400}$ (650)
*Nru*I	450	700 1100	—	—	—
*Sfi*I + *Mlu*I	200	(500) $\underline{650}$	$\underline{220}$ 350	$\underline{230}$ (150)	$\underline{300}$ (500) $\underline{540}$ 680
*Sfi*I + *Bss*HII	$\underline{200}$	400 520 580 650	$\underline{220}$ 350	$\underline{230}$	390 680 800
*Sfi*I + *Sac*II	$\underline{200}$	500 650	$\underline{220}$ 350	$\underline{230}$	390 680 800
SfiI + *Sal*I	ND	ND	$\underline{220}$	$\underline{220}$ (450)	$\underline{200}$ (450)
*Sfi*I + *Nru*I	200	$\underline{420}$ $\underline{650}$	$\underline{220}$ 350	$\underline{230}$	$\underline{300}$ 390 680 800
*Sac*II + *Bss*HII	ND	400 700 900	ND	ND	(710)
*Sac*II + *Nru*I	ND	$\underline{250}$ 450	ND	ND	ND
*Sac*II + *Mlu*I	ND	(300) 700	ND	ND	ND
*Sal*I + *Nru*I	ND	ND	ND	$\underline{400}$ (650)	$\underline{400}$ (650)
*Sal*I + *Mlu*I	ND	ND	300	(120) $\underline{400}$ 630	(320) $\underline{400}$ (480) 630
*Mlu*I + *Nru*I	ND	ND	ND	630 700 850 900 >950	(360) 730 700 850 900 >950
*Nru*I + *Bss*HII	ND	350 400 750	ND	ND	ND

Sizes of bands are in kilobase pairs. A dash indicates that no band was observed; ND, not determined. To confirm identity of bands recognized by two probes, several gels were run using different pulse times giving optimal resolution for different size ranges, minimizing the accidental comigration of non identical bands. For the construction of the map, we also considered the intensity of bands relative to the strong *Sfi*I-bands. Strong bands (>30% of the intensity of *Sfi*I, considered cut to completion) are underlined and very weak bands (<5% the intensity of *Sfi*I) are in brackets. Note that weak bands can arise from a site which is cut to completion, when the second site is incompletely cut. The partial bands produced by *Sfi*I as recognised by pXJ1.1 and L1-4 might reflect methylation in recognition sequences of the type GGCmCGNNNNNGGCC or GGCCNNNNNGGCmCG. In this case observation of two bands due to an RFLP can be excluded, since two bands are also observed in an *Sfi*I digest of DNA from male blood cells (not shown). *Nae*I, although often producing bands on pulsed-field gels, gave a rather complex pattern with four and more bands (see Fig. 1*b*). These could not be interpreted without ambiguities. *Nae*I-fragments recognized by probes were therefore not considered in the construction of the map, and are not listed here.

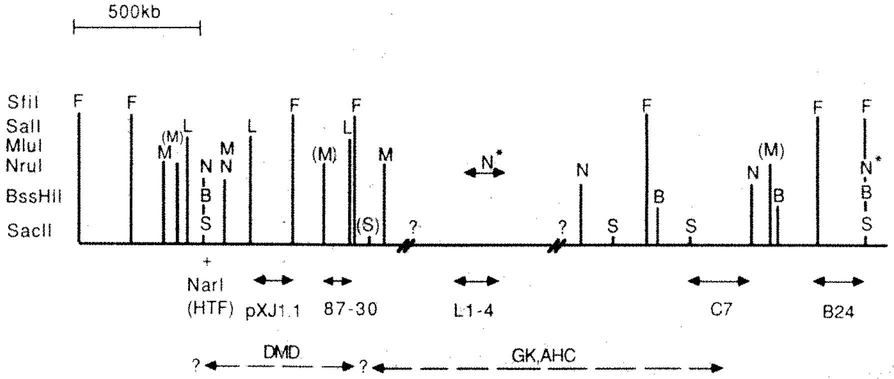


Fig. 2 Long range restriction map. The data for producing the map are summarized in Table 1. Weakly cut (highly methylated) sites are in brackets. The approximate locations of the DMD-gene(s) and of the glycerol kinase (GK) and congenital adrenal hypoplasia (AHC) genes are indicated below the map. Note that physical linkage of probes is only considered proven when at least three common bands generated by different enzymes were observed. Limits of localization for the probes used are indicated by arrows around the probe names below the map. In addition to the sizes of bands detected, the methylation status of a site was also considered for the construction of the map. The degree of methylation was estimated from the ratio of intensity of bands in double digests of enzymes cutting completely with the enzyme in question. Sizes were estimated with respect to yeast and λ size markers. It should be noted that the absolute sizes can only be considered to be approximate, because the amount of DNA loaded and pulse times[19] can slightly change the apparent size of a fragment. Comigration of two fragments was considered to be reliable, since the same blots were used for different probes. Since the cell line LCL127 is not homozygous for the X-chromosome, we can not rule out that some partial cutting is due to polymorphisms rather than methylation. Thus the map might have combined sites from two different chromosomes in a few instances. For the pERT87-30/pXJ1.1 cluster (Fig. 1*a*) subclones pERT87-30, localized about 50 kb distal of pERT87-15 (ref. 3 and Monaco and L. Kunkel, personal communication) and pXJ1.1 were used. Both pERT87-30 and pXJ1.1 recognize the same *Sal*I-fragment, but different *Sfi*I-fragments. Double digests of *Sfi*I and *Sal*I localize the *Sfi*I-site between the two probes approximately to the middle of the *Sal*I-fragment. Therefore, probe pXJ1.1 is located within 200 kb of the end of the large *Sfi*I fragment. Double digests with *Sfi*I/*Mlu*I and *Sfi*I/*Nru*I give the position of *Mlu*I and *Nru*I sites from the end of the *Sfi*I-fragment, similar to partial cosmid mapping[23]. A very weakly cutting (highly methylated) *Mlu*I-site within the *Sal*I fragment is detected in *Sal*I/*Mlu*I double digests. This *Mlu*I site cleaves the *Sal*I fragment into fragments of about 100 kb (hybridizing to pERT87-30) and of about 300 kb (hybridizing to pXJ1.1). The subclone pERT87-1, about 100 kb closer to pXJ1.1 than pERT87-30, recognizes the 300-kb *Sal*I-*Mlu*I fragment (not shown). This *Mlu*I-site is localized within the small *Sfi*I-fragment seen by pERT87-30, since it is also detected in a *Sfi*I/*Mlu*I digest by pERT87-30 as a weak *Sfi*I/*Mlu*I band of ~100 kb. Therefore the distance between pERT87-30 and pXJ1.1 is at least 100 kb, the distance between the weakly cut *Mlu*I site and the *Sfi*I-site separating the two probes. In the B24/C7 cluster (Fig. 1*b*), Probes B24 and C7 recognize the same *Sac*II fragments, but different *Sfi*I-fragments. A *Sac*II/*Sfi*I double digest then identifies a *Sac*II-site in the *Sfi*I-fragment of C7. The size of the common *Sac*II fragment predicts that the *Sfi*I fragments recognized by B24 and C7 are adjacent. There are different *Nru*I fragments recognized by these two probes. A *Nru*I/*Sfi*I double digest localizes one *Nru*I-site within the *Sfi*I-fragment seen by C7. Since the *Nru*I fragments are not shared, they point in opposite direction from this *Nru*I-site. This state of the map predicted a 250-kb *Nru*I/*Sac*II fragment recognized by C7, which could be confirmed (data not shown). Similarly, the *Bss*HII-sites were localized. Since *Nru*I does not cut the *Sfi*I-fragment seen by B24, the *Nru*I sites leading to a 450-kb *Nru*I fragment around B24 cannot be localized except for flanking the *Sfi*I fragment. A suggested *Nru*I-site (asterisk) is shown at the right hand side of the map. The additional *Nru*I-fragment of about 1,100 kb hybridizing to C7 seems not to be corecognized by B24, thus would be localized further to the left of the map (proximally). But as B24 hybridizes rather weakly, we can not completely rule out that the hybridization signal in this case was too weak to be detected. Therefore, the localization of the *Nru*I-site in the direction of L1-4 (asterisk) is considered not proven. The probe L1-4 could not link to any of the clusters (not shown). A 850-kb *Bss*HII fragment is recognized, and two *Sfi*I-fragments of 220 and 350 kb. As no bands are shared, the minimal distance between the *Mlu*I site distal of pERT87-30 and the *Nru*I site proximal to C7, is 500 kb, giving the distance of C7 to pERT87-30 as at least 1,000 kb.

recognized by L1-4 and any of the other probes. Taking the largest fragments not shared into account, the distance between the probes C7 and pERT87-30 must be at least 1,000 kb (Fig. 2).

When hybridizing pXJ1.1 to DNA restricted with *Sfi*I and several CpG-recognizing enzymes (*Nru*I, *Nar*I, *Sac*II and *Bss*HII) new fragments of identical size (390 kb) were detected in each case (see Fig. 1*c*). Since the probe pXJ1.1 is localized close to one end of the *Sfi*I fragment, these four sites must be localized, within the resolution of the gel, in close proximity to each other, 200-300 kb proximal to the probe pXJ1.1 (see map in Fig. 2). Such a clustering of sites would be expected in the case of an HTF (*Hpa*II tiny fragments) island[14,15], often found in close proximity to the 5′ end of genes. Very recently (data not shown) we found that at least some sites in the postulated HTF-island are cut to completion in DNA isolated from human sperm. The partial cutting of the sites observed here (Fig. 1*c*) is probably due to methylation of HTF-islands in inactivated X chromosomes[14]. Thus, with respect to methylation, this cluster behaves like known HTF-islands on the X-chromosome[14]. Since deletions in the pXJ1.1 region have a DMD phenotype as their only common effect[16,17], such an CpG-rich island could indicate the promoter sequence for the gene responsible for the DMD mutation. This would require the DMD gene to have a size of at least 600 kb (distance of the HTF island to the start point of the most distal published deletion[3]), with transcription proceeding in the direction centromere to telomere. Alternative models (for example, multiple clustered genes responsible for the same phenotype) can not be ruled out, but such a large gene size would be compatible with both the high incidence of DMD mutations and the observed distribution of deletions, translocations and other mutations associated with the disease throughout a very large chromosomal segment.

A complicating observation in the PFG analysis was partial cleavage of sites, resulting in multiple and weak bands recognized by these unique probes (compare, for example, the *Sfi*I-bands to all others in Fig. 1*a*, *b*). Partial restriction is caused by the sensitivity of all enzymes cutting rarely in the mammalian genome to CpG methylation in their recognition site. On the other hand, these partial digests added important information for mapping the region, since more sites can be analysed. Some of these sites might be due to polymorphisms since the lymphoblastoid cell line used contains four X chromosomes. This cell line was found to yield more visible bands than, for example male blood, as a result of a higher signal and/or hypomethylation. Nevertheless, major parts of the maps could be confirmed using DNA isolated from blood or sperm (not shown). All rare-cutting enzymes (except *Sfi*I, but see Fig. 1 legend) contain at least one, and usually two, CpG sequences in their recognition site. Reduced cleavage by CpG-recognizing enzymes is especially noticeable in the region surrounding the DMD locus, which appears to be depleted of unmethylated CpG-rich sites compared to other regions of the human genome which have been examined in our laboratory (M. Búcan *et al.*, unpublished data). For example, the enzyme *Not*I does not cleave within the entire region analysed. This obvious sparsity of unmethylated CpG-rich sequences might be correlated with a relatively low density of genes in this region, also indicated by the fact that large deletions around the DMD locus are associated with very few phenotypes[16,17]

We have described here partial physical maps covering ~3000 kb, about 2% of the human X chromosome. Such a map will be extremely useful in determining the exact size of the DMD gene(s), and will allow us to estimate the positions and extents of translocations and deletions in this area. The immediate positioning of new clones in the DMD area is now possible, with a resolution difficult to reach by purely genetic approaches.

We are especially grateful to Denise P. Barlow, who developed the PFG technology. We thank Anthony Monaco and Louis Kunkel, Peter Ray and Ron Worton, and Jean-Louis Mandel for providing the probes, Uta Francke for introducing us to the DMD problem, Marija Bucan for preparing some of the filters used, Gert-Jan van Ommen and Louis Kunkel for discussions, Denise P. Barlow, Michael Hortsch, Frank Michiels and Lisa Stubbs for critical reading of the manuscript, and Carol Stettner and other members of the electrical and mechanical workshops at the EMBL, for building the pulsed field gradient gel equipment.

Note added in proof: The recent publication of a cDNA from the DMD gene (A. M. Monaco *et al.*, *Nature* **323**, 646-650) confirms the direction of transcription and minimal gene size postulated here.

Received 28 July; accepted 29 October 1986.

1. Kedes, L. H. *Trends Genet.* **1**, 205-209 (1985).
2. Boyd, Y. & Buckle, V. J. *Clin. Genet.* **29**, 108-115 (1986).
3. Kunkel, L. M. *et al. Nature* **322**, 73-77 (1986).
4. Schwartz, D. C. & Cantor, C. R. *Cell* **37**, 67-75 (1984).
5. Smith, C. L. *et al.* in *Genetic Engineering* (eds J. K. Setlow & A. Hollander) Vol. 8, 45-70 (Plenum, New York, in the press).
6. Monaco, A. P. *et al. Nature* **316**, 842-845 (1985).
7. Ray, P. N. *et al. Nature* **318**, 672-675 (1985).
8. De Martinville, B. *et al. Am. J. hum. Genet.* **37**, 235-249 (1985).
9. Goodfellow, P. N., Davies, K. E., Ropers, H.-H. *HGM* **8**, *Cytogen. Cell genet.* **40**, 296-352 (1986).
10. Kunkel, L. M. *et al. Gene* **33**, 251-258 (1985).
11. van Ommen, G. J. B. *et al.* submitted.
12. Bertelson, K. *et al. J. med. Genet.* (in the press).
13. Dunger, D. B. *et al. Lancet* **i**, 585-587 (1986).
14. Bird, A. P. *Nature* **321**, 209-213 (1986).
15. Brown, W. A. R. & Bird, A. P. *Nature* **322**, 477-481 (1986).
16. Francke, U. *et al. Am. J. hum. Genet.* **37**, 250-267 (1985).
17. Wilcox, D. E. *et al. Hum. Genet.* **73**, 175-180 (1986).
18. Frischauf, A.-M. *et al. J. molec. Biol.* **170**, 827-842 (1983).
19. Bernards, A. *et al. Gene* **42**, 313-322 (1986).
20. Herrman, B. G. *et al. Cell* (in the press).
21. Feinberg, A. P. & Vogelstein, B., *Analyt. Biochem.* **132**, 6-13 (1983).
22. Herrmann, B. G. *et al. Cell* 469-476 (1986).
23. Rackwitz, H. R. *et al. Gene* **40**, 259-266 (1985).

Recombination between immunoglobulin variable region gene segments is enhanced by transcription

T. Keith Blackwell, Mark W. Moore*, George D. Yancopoulos, Heikyung Suh, Stuart Lutzker, Erik Selsing* & Frederick W. Alt

Department of Biochemistry and Microbiology, Institute for Cancer Research, Columbia University College of Physicians and Surgeons, 630 W. 168th Street, New York, New York 10032, USA
* Rosensteil Basic Medical Sciences Research Center and Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA

Immunoglobulin (Ig) variable (V) region genes are assembled in precursor B (pre-B) lymphocytes from multiple germline segments. The heavy-chain V-region gene is composed of variable (V_H), diversity (D) and joining (J_H) segments; kappa (κ) and lambda (λ) light-chain V-region genes have analogous V_L and J_L segments. Assembly of Ig V-gene segments, as well as those of the highly related T-cell receptor, is regulated at several levels and shows both stage and tissue specificity; for example Ig heavy-chain V-gene assembly precedes that of Ig light chains during B-cell differentiation[1]. Joining of all classes of V-gene segments involves conserved recognition sequences that are probably targets for a common recombinase[2]. Evidence has been presented suggesting that rearrangement of specific classes of segments is regulated by modulation of their accessibility to the recombinase[3]. To elucidate mechanisms which control V-region gene assembly, we have investigated the effect of flanking gene expression on the frequency at which introduced V-gene segments are assembled in pre-B cell lines. Our findings suggest that transcription may play a direct role in the regulation of immunoglobulin V-gene assembly.

Transformation by Abelson murine leukaemia virus (A-MuLV) generates pre-B cell lines; during culture some lines

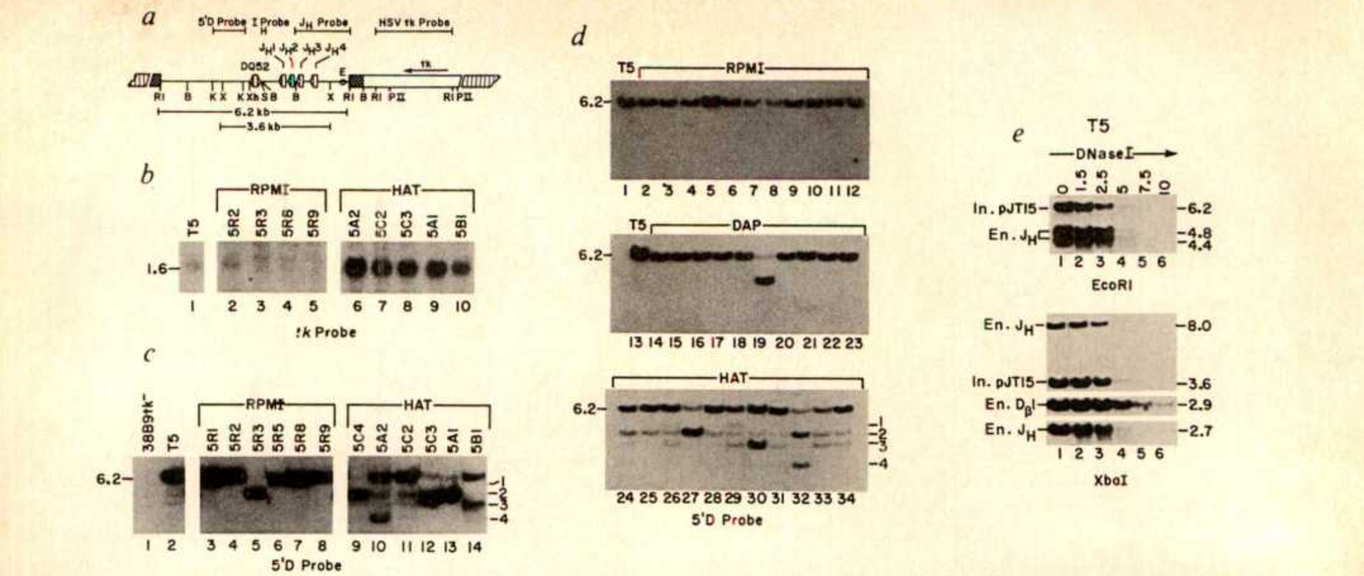


Fig. 1 Activation of transfected heavy-chain *V*-gene assembly. a, Map of the T5 integration. Line, murine DNA; checked box, plasmid DNA; open box, *tk* DNA; and lined box, cellular sequences. *D* and J_H segments with associated recognition sequences (triangles) are indicated. E, the heavy chain enhancer; arrow, *tk* gene. Restriction endonuclease sites are as follows: RI, *Eco*RI; B, *Bam*HI; K, *Kpn*I; X, *Xba*I; Xh, XhoI; S, *Sac*I; PII, *Pvu*II. *b*, *tk* expression in T5 derivatives. Total RNA (10 μg) from the cell lines indicated was assayed by Northern blotting for hybridization to a ^{32}P-labelled HSV *tk* probe as described[3]. *c*, Rearrangement of the integrated substrate. Genomic DNA (10 μg) from cell lines indicated was digested with *Eco*RI and analysed by Southern blotting for hybridization to a ^{32}P-labelled 5'D probe as described[4]. *DQ*52 to J_H 1-4 rearrangements are indicated by 1-4. *d*, Substrate rearrangements in T5 subclones isolated in RPMI (lanes 2-12), DAP (lanes 14-23) and HAT (lanes 23-34). *e*, DNase sensitivity of the integrated substrate. Nuclei isolated from T5 were incubated with 0-10 μg ml^{-1} of DNase I as described[2]. Genomic DNA (10 μg) was digested with either *Eco*RI or *Xba*I, hybridized to the J_H and D_B1 (ref. 2) probes, and assayed as in *c*. Endogenous (En.) and substrate (In.) fragments are indicated. Densitometry revealed that in lane 4 the pJT15 substrate and the 2.7-kb endogenous J_H-containing fragment retain 6 and 10%, respectively, of their lane 1 signal intensity, whereas the endogenous D_B1-containing fragment retains 81%. Fragment sizes are in kb.

Methods. The 38B9*tk*⁻ cell line has been described[4]. T5 was derived by CaPO$_4$ co-transfection of 5×10^7 38B9 *tk*⁻ cells with pJT15 and pSV2neo, followed by selection in G418 (ref. 4). HAT- or DAP-resistant clonal T5 derivatives were isolated in 24-well plates by plating 10^4-10^5 cells per well in 5×10^{-5} M DAP or in HAT[4]. For DNase sensitivity assays, T5 and 5R9 were expanded in RPMI and 5B1 in HAT.

assemble endogenous immunoglobulin *V*-region genes[1]. The Abelson line 38B9*tk*⁻ has been shown to assemble endogenous and transfected heavy-chain *V*-gene segments[4]. To examine the effect of local gene expression on rearrangement activity, a substrate (pJT15) containing *DQ*52 and J_H segments plus the heavy-chain enhancer downstream from a herpes simplex virus thymidine kinase (*tk*) gene was introduced into 38B9*tk*⁻ by unlinked co-transformation[4]. We used transfected clone T5, which had integrated a single copy of pJT15 (Fig. 1). As observed in 38B9 cells transfected with other *tk*-based constructs[4], *tk* messenger RNA was barely detectable in T5 cells (Fig. 1*b*, lane 1; see below). To isolate rare cells that expressed the *tk* gene, T5 populations were plated in HAT (hypoxanthine, aminopterin and thymidine) medium[4]; HAT-resistant clones occurred at a frequency of 10^{-4} to 10^{-5}, and all expressed normal 1.6 kb (kilobase) *tk* RNA from the *tk* promoter (Fig. 1*b*, lanes 6-10). Clones isolated in non-selective medium did not appear to express *tk* mRNA (Fig. 1*b*, lanes 2-5). The ability to isolate *tk*-expressing variants from transfected populations formed the basis of our experiments.

The region 5' of *DQ*52 (5'D probe, Fig. 1*a*) has been deleted from the 38B9 genome through endogenous joining[4] (Fig. 1*c*, lane 1). Thus, in T5 derivatives, the unrearranged pJT15 substrate was detected as a 5'D-positive, 6.2-kb *Eco*RI fragment (Fig. 1*c*, lane 2). In a given T5 cell, the *DQ*52 segment may join to any of the four J_H segments in the substrate; therefore, rearrangements which occurred during expansion of T5 subclones were detected as a mixture of four smaller, rearrangement-specific 5'D-positive *Eco*RI fragments (Fig. 1*c* and 1*d*). Five such rearranged fragments were molecularly cloned; sequence analysis indicated that all contained normal DJ_H joins (Fig. 2), some of which contained N regions[5]. Insertion of N regions into DJ_H rearrangements within the pJT15 construct correlated with terminal deoxynucleotidyl transferase RNA levels in clonal derivatives of a single line (Fig. 2), providing the strongest evidence to date linking this enzyme with N-region addition[2,5].

Genomic DNA from the initial T5 population contained a very low but detectable level of specific rearrangements (Fig.

```
DQ52                 CAACTGGGACCACGGTG
                                 *
1 (+) 5A2-1(DJH4)    CAACTGGG---TGAG---GCTATGGAC
2 (+) 5B1-1(DJH4)    CAACTGGGA----ACTATGCTATGGAC
JH4                      TATTGTGATTACTATGCTATGGAC

                           *
3 (+) 5B1-2(DJH3)    CAAC---CC--GCCTGGTTTGCTTACT
JH3                      CACAATGTGCCTGGTTTGCTTACT

4 (-) I12-1(DJH2)    CAACTGGG-------------CTACTGG
5 (-) I12-2(DJH2)    CAACTGG---------TTTGACTACTGG
JH2                     TAGTGTGACTACTTTGACTACTGG
```

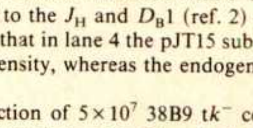


Fig. 2 Nucleotide sequences of *DQ*52 to J_H joins. *Eco*RI fragments hybridizing to the 5'D probe but not the I probe (Fig. 1*a*) were isolated from Charon 16A libraries of 5A2, 5B1 and I12 DNA as described[4]. Joins isolated from terminal deoxyribosyl transferase (TdT)-positive (+) and TdT-negative (−) lines are noted; nucleotide sequences were determined as described[4]. Recognition heptamers are underlined, and joins aligned with germline sequences so that bases lost during joining are apparent. Asterisks, N regions. Lower right panel: Total RNA from the lines indicated was assayed for hybridization to a TdT cDNA clone as described[2]. The TdT transcript and its approximate size in kb are indicated.

Table 1 Substrate rearrangement correlation with HAT resistance

	Experiment no. 1		Experiment No. 2		Experiment No. 3		Experiment No. 4		Total	
	RPMI	HAT	RPMI	HAT	RPMI	HAT	RPMI	HAT	RPMI	HAT
Subclones examined	7	5	6	9	11	11	11	11	35	36
Unrearranged	4	0	5	0	10	2	10	0	29	2
Single rearrangement	2	0	1	0	1	1	0	0	4	1
Rearranged*	1	5	0	9	0	8	1	11	2	33

*DQ*52 to J_H joining was assayed in clonal T5 derivatives in four separate experiments.

* A level of *DQ*52 to J_H joining within the population at or greater than that of Fig. 1*d*, lane 4.

1*c*, lane 2), suggesting the presence of a minor subpopulation that had joined introduced *D* and J_H segments. Clonal populations derived in RPMI did not express the *tk* gene or show evidence of substrate rearrangement (Fig. 1*c* and *d*); a few tk^- clones had a single *D* to J_H join (Fig. 1*c*, lane 5) demonstrating that *D* to J_H rearrangement *per se* did not initiate expression of the contiguous *tk* gene (Fig. 1*b*, lane 3). Such joins probably occurred before subcloning, as cells containing the 6.2-kb unrearranged substrate were absent. In contrast, *tk*-expressing clonal populations isolated in HAT medium displayed a striking level of rearrangement within the substrate (Fig. 1*c* and *d*). Most HAT-resistant clonal populations also contained many cells with an unrearranged 6.2-kb substrate, indicating each cell derived from a progenitor that expressed an unrearranged *tk* substrate, and that joining occurred during expansion of the population (Fig. 1*c* and *d*). Subcloning of initial HAT-resistant clones confirmed that *D* to J_H joining within the substrate occurred frequently in *tk*-expressing derivatives (not shown). HAT resistance (*tk* expression) was correlated with rearrangement of the introduced *D* and J_H segments in four separate experiments (Table 1).

To investigate whether increased recombination frequencies in HAT-resistant cells were associated with *tk* expression rather than other potential effects of selection (such as interference with nucleotide metabolism), T5 cells were also grown in 2,6-diaminopurine (DAP) which selects against cells expressing adenosine phosphoribosyl transferase[6]; conditions were employed such that survival frequency in DAP was less than in HAT. Surviving clones (Fig. 1*d*, lanes 14–23) showed low levels of rearrangement, as did those derived in RPMI (Fig. 1*d*, lanes 2–12). Therefore, the increased frequency of *DQ*52 to J_H rearrangement in HAT-resistant clones is associated specifically with expression of the flanking *tk* gene.

The 'active' chromatin state required for expression of eukaryotic genes is reflected by its sensitivity in isolated nuclei to digestion with DNase I (DNase sensitivity)[7,8]. To define the requirements for *V*-gene assembly further, we examined relative DNase sensitivity of the pJT15 construct in the nuclei of the T5 line, the HAT-derived 5B1 line (which expresses the *tk* gene and actively rearranges the substrate) and the RPMI-derived 5R9 line (which does not express the *tk* gene or rearrange the substrate). In all three the DNase sensitivity was similar to that of the active endogenous J_H loci; all were DNase sensitive relative to the inactive[2] D_B1 locus of the T-cell receptor (Fig. 1*e*, lanes 1–6; only T5 data are shown). Thus, DNase sensitivity *per se* is not responsible for active rearrangement of the introduced construct.

To analyse further the basis for low frequency expression of introduced HSV *tk* genes in 38B9 and the relationship between increased *tk* expression and recombination in HAT-resistant derivatives, we analysed a 38B9 *tk*-transformant, L1, which contained multiple intact integrations of a construct containing $V_\lambda 1$ and $J_\lambda 1$ elements flanking a *tk* gene (pLT2, Fig. 3*a*). In this case, transformation was accomplished by unlinked co-electroporation[2]. Assuming that low frequency expression of a co-transfected *tk* gene in 38B9tk^- derivatives resulted from a '*cis*-acting' transcriptional modulation, it would be expected that HAT selection would isolate a population of cells which expressed the *tk* gene in only one of the integrated substrate copies. Because expressed and non-expressed construct copies can be distinguished (see below), we could directly compare the effect of differential expression on rearrangement activity of identical constructs within single cells. Although the 38B9 line is in the heavy-chain phase of its *V*-region gene assembly programme and does not rearrange endogenous λ light chain genes at a detectable level, earlier evidence suggests that all immunoglobulin gene segments are rearranged by a common recombinase[2]; therefore, introduced λ gene segments, if accessible, should be rearranged in this line. Integrated pLT2 copies were detected in genomic DNA of L1 derivatives as $V_\lambda 1$-positive *Pvu*II fragments of different sizes; a *V–J* rearrangement (for example VJ1, Fig. 3*c*) was detected as a $V_\lambda 1$-, $J_\lambda 1$-positive *Pvu*II fragment 1.7 kb smaller than the $V_\lambda 1$-positive *Pvu*II fragment containing the copy from which it was derived (for example RC1, Fig. 3*c*). Nucleotide sequence analysis of a *VJ*1 join molecularly cloned from L1HB confirmed that it occurred by the normal mechanism (Fig. 3*d*).

A population of L1 cells that expressed the introduced *tk* gene was isolated in HAT (L1H, Fig. 3*b*); these HAT-resistant cells again appeared at a frequency of 10^{-4} to 10^{-5}. With respect to integrated construct copies, the genomic DNA of the L1H population appeared identical to that of L1 except for complete loss of a substrate copy which was identified by a 4.5 kb $V_\lambda 1$-positive *Pvu*II fragment (see Fig. 3*c*). Assuming a common recombinase[2], deletion of the expressed *tk* gene(s) as a result of V_λ to J_λ joining within the construct would be expected to be the major mechanism of BUdR(bromodeoxyuridine)-resistance in these lines[4]. Removal of L1H from HAT (L1HT, Fig. 3*b*) followed by selection in BUdR yielded the tk^- population L1HB. These cells had completely rearranged a single integrated construct (RC1) to its product *VJ* join (*VJ*1) (Fig. 3*c*, lanes 3 and 7), whereas other full length copies remained intact. This identifies RC1 as the *tk* copy expressed in this subclone and confirms that the *tk* mRNA found in HAT-resistant subclones derives from a single integrated substrate. This rules out potential post-transcriptional (*trans*-acting) mechanisms for increased *tk* expression that would affect expression from all copies, and strongly suggests that low frequency expression of transfected *tk* genes in these lines is mediated at the transcriptional level.

To identify independently pLT2 construct loci that are expressed, we measured relative DNase sensitivity of the integrated constructs. L1H was expanded in non-selective medium, generating population L1HT2 (Fig. 3*b*), in which RC1 undergoes rearrangement (Fig. 3*e*, lane 1; see below). In L1HT2, both RC1 and its rearranged product *VJ*1 are clearly DNase sensitive relative to endogenous genes and non-rearranging pLT2 integrations (Fig. 3*e*, lanes 1–6). This provides additional evidence that *tk* expression occurs primarily from RC1 in L1H and that activation of this copy is at the transcriptional level. Significantly, after a brief passage in non-selective medium, L1H cells converted the RC1 integration to its derivative rearrangement; however, no rearrangement of other integrated copies was observed (Fig. 3*c*, lanes 4 and 8). Thus, the transcribed construct copy has a high rate of rearrangement but non-expressed copies in the same line do not; this is not because of defects in their

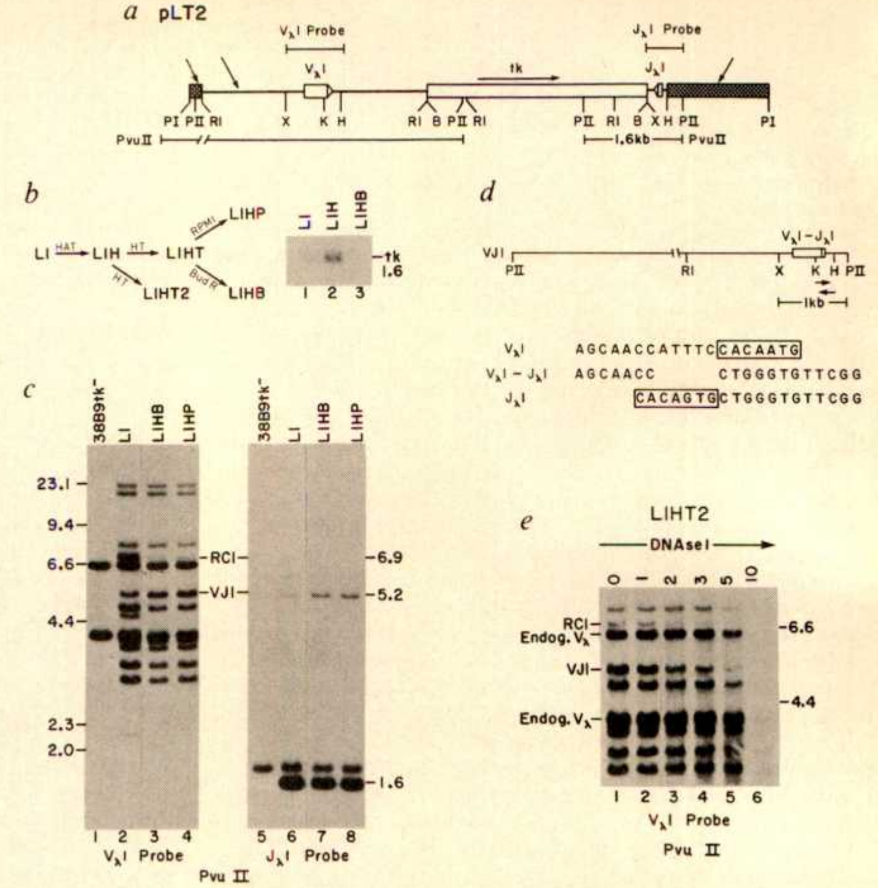


Fig. 3 Rearrangement of transcribed light chain substrates. *a*, Map of the pLT2 substrate. Structures and restriction endonuclease sites as in Fig. 1*a* (also H, *Hin*dIII; PI, *Pvu*I). Arrows above map, regions within which pLT2 copies integrated into cellular DNA. *b*, Selection of L1 derivatives and *tk* expression. Total RNA (10 μg) was assayed for *tk* expression as in Fig. 1*b*. *c*, Rearrangement of the pLT2 substrate. Genomic DNA (10 μg) was digested with *Pvu*II and analysed as in Fig. 1*c* for hybridization to the probes indicated. Sizes are in kb. *d*, Restriction map and nucleotide sequence of a *VJ*1 rearrangement. To isolate this join, L1HB genomic DNA was digested with *Pvu*II, *Eco*RI linkers added and fragments cloned into Charon 16A and screened with $V_\lambda 1$ and $J_\lambda 1$ probes. The join is aligned with germline sequences so that the bases lost are apparent. *e*, DNase sensitivity of integrated and endogenous (Endog.) λ loci. Nuclei from L1HT2 were digested with 0–10 μg ml^{-1} of DNase I; genomic DNA was then digested with *Pvu*II and assayed for hybridization to the $V_\lambda 1$ probe as described in Fig. 1*c*. Fragments corresponding to RC1, *VJ*1 and the endogenous V_λ loci are indicated. Densitometry revealed that in lane 4, RC1 and *VJ*1 retained 17 and 36%, respectively, of their lane 1 signal intensity. In lane 4 the pLT2 copy which is slightly larger than RC1, and the endogenous 6.4-kb V_λ-containing fragment retained 87 and 97%, respectively, of their original intensities. Sizes are in kb.

Methods. pLT2 was derived from pUC13. pLTRneo contains the TN*5neo* resistance gene transcribed from a Moloney-MuLV LTR promoter. L1 was derived by electroporation[2] of 7.5×10^6 38B9 tk^- cells with 10 μg of pLTRneo and 40 μg of pLT2 which had been linearized with *Pvu*I. Transformants were isolated in G418 as in Fig. 1*a*. L1 was passaged through HAT, HT and 10^{-4} M BUdR as described[4]. Other analyses were as described in Figs 1 and 2.

structure, as some different copies were expressed and rearranged in independent L1 derivatives (not shown).

Identical *tk*-based recombination substrates are expressed frequently (>25%) when introduced into myeloma (J558) or L cell lines by the same procedures (not shown) that lead to infrequent expression (10^{-4}–10^{-5}) in pre-B lines. This may be related to the relatively low transformation frequency of pre-B lines[4]. Current results suggest that the block in pre-B cells is transcriptional but does not result from lack of general DNase sensitivity (Fig. 1) or obvious methylation-related phenomena (not shown). Significantly, we found no site-specific recombination events in L cell or J558 transformants when they were subjected to analyses identical to those described above (data not shown), suggesting that recombinase activity may be regulated with respect to stage and tissue specificity. The assay would have detected recombinase activity 3–4 orders of magnitude lower than found in several pre-B cell lines. However, many more different lines must be assayed to verify this point.

In T5 derivatives, the pJT15 substrate is of equivalent DNase-sensitivity to the endogenous J_H loci regardless of *tk* gene expression (Fig. 1*e*); this sensitivity is probably conferred by the associated heavy-chain enhancer[9–10] which is active in pre-B cells[11]. Therefore, activity of the enhancer element (and DNase-sensitive structure *per se*) on the pJT15 substrate appears insufficient to confer direct activation of *D* to J_H rearrangement. T-cell receptor *V*-gene segments introduced into 38B9 within an *expressed tk*-based recombination substrate were DNase sensitive and had a high spontaneous rearrangement rate[2]. In the λ substrate which did not contain known enhancer elements, DNase sensitivity (and rearrangement) appeared to occur only in association with transcription of the *tk* gene. Apparently, efficient assembly of *V* segments requires either transcription or factors closely associated with transcription. Our results do not preclude a regulatory role for enhancer elements with respect to rearrangement; they may allow transcription or closely-associated events which then promote rearrangement. The lack of a clear function for known or potential products of germline immunoglobulin or T-cell receptor gene transcripts supports a direct role for transcription[1]. A relationship between local gene expression and recombination has been observed in the context of other genomic rearrangement events, such as the yeast mating-type locus switch[12] and the immunoglobulin class switch [13,14]. Transcription could influence accessibility for rearrangement in several ways. The recombination mechanism might involve transcription itself; for example DNA unwinding by RNA polymerase[15] could be required for initiation of recombination. Alternatively, transcription of these loci might be associated with structural or nuclear locational changes in DNA that are not detected by general DNase sensitivity and not solely mediated by the enhancer[8,16].

We thank Y. Rosenberg, B. Malynn and P. Ferrier for critically reading this manuscript, and C. S. H. Young, S. Morrison and S. Silverstein for helpful discussions. Supported by NIH Grants AI-20047, CA-40427 and CA-2112-09, and a Searle Scholars Award to F.W.A. (who is an Irma T. Hirschl Career Scientist and a Malinckrodt Scholar) and by NIH grant CA-36370 and NSF grant DCB-8302110 to E.W.S. T.K.B., S.L. and G.D.Y. were supported by NIH training grant GM-07367, and M.W.M. by NIH training grant AI-07069.

Received 26 June; accepted 23 September 1986.

1. Alt, F. W., Blackwell, T. K., De Pinho, R.A., Reth, M. G. & Yancopoulos, G. D. *Immun. Rev.* **89**, 5–30 (1986).
2. Yancopoulos, G. D., Blackwell, T. K., Suh, H., Hood, L. & Alt, F. W. *Cell* **44**, 251–259 (1986).
3. Yancopoulos, G. D. & Alt, F. W. *Cell* **40**, 271–281 (1985).
4. Blackwell, T. K. & Alt, F. W. *Cell* **37**, 105–112 (1984).
5. Alt, F. W. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4118–4122 (1982).

6. Jones, G. E. & Sargent, P. A. *Cell* **2,** 43-54 (1974).
7. Weintraub, H. & Groudine, M. *Science* **193,** 848-856 (1976).
8. Weintraub, H. *Cell* **42,** 705-711 (1986).
9. Banerji, J., Olson, L. & Schaffner, W. *Cell* **33,** 729-740 (1983).
10. Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33,** 717-728 (1983).
11. Gerster, T., Picard, D. & Schaffner, W. *Cell* **45,** 45-32 (1986).
12. Klar, A. J. S., Strathern, J. N. & Abraham, J. A. *Cold Spring Harb. Symp. quant. Biol.* **49,** 77-88 (1984).
13. Stavnezer-Nordgren, J. & Sirlin, S. *EMBO J.* **5,** 95-102 (1986).
14. Yancopoulos, *et al. EMBO J.* (in the press).
15. Fisher, M. L. *Nature* **299,** 105-106 (1982).
16. Hutchinson, N. & Weintraub, H. *Cell* **43,** 471-482 (1985).

Effect of mutations at the lariat branch acceptor site on β-globin pre-mRNA splicing *in vitro*

H. Hornig, M. Aebi & C. Weissmann

Institut für Molekularbiologie I, Universität Zürich,
8093 Zürich, Switzerland

Introns are excised from full-length transcripts (pre-messenger RNAs) of eukaryotic genes in two steps[1,2]. First, the pre-mRNA is cleaved at the 5′ splice site and a branched (lariat) intermediate is formed[1-6]. Then, cleavage at the 3′ splice site and ligation of the two exons leads to the release of the lariat intron. The intron sequence which accepts the 5′ end to form the lariat branch is strictly conserved in yeast[7-9], but shows more variation in eukaryotes[10,11]. To investigate the requirements for branch formation in eukaryotes further, we have studied *in vitro* splicing of a rabbit globin gene intron with mutations of the normal branch-accepting adenosine nucleotide. We conclude that all four nucleotides can serve as branch acceptors, but that A and C are preferred to G and U in lariat formation. Mutation of the normal A to G or U can lead to an A residue one nucleotide upstream of the normal branch site being used instead. Only branches to A or C participate efficiently in the second splicing step.

Lariat formation[1-5] results from the 2-5′ linkage of the 5′ end of the intron to an adenosine residue close to the 3′ splice site[6]. In the case of yeast nuclear pre-mRNA splicing, efficient lariat formation depends stringently on the strictly conserved TACTAAC-box, where the underlined 3′-proximal A residue serves as branch acceptor[4,5]. Deletions or point mutations within this sequence abolish splicing[7-9]. In contrast, the branch acceptor sequence of higher eukaryotes is poorly conserved[10,11] and deletions of this sequence lead to the activation of cryptic branch points which allow normal splicing, albeit at a reduced rate[12,13]. There seems to be a distance constraint as regards branch formation, in that the acceptor sequences identified so far lie within 18-37 nucleotides upstream of the 3′ splice site[11,12].

The branch acceptor nucleotide of the large rabbit β-globin intron has been determined *in vivo*[3] as well as *in vitro* [13] to be predominantly the underlined A residue and to a lesser extent the A residue preceding it, in the sequence GCTAAC. In order to study the nucleotide requirement at the branch point, we have converted the principal branch acceptor A residue (position 1,036)[14] to either G, C or T by site-directed mutagenesis. The mutated β-globin sequences (*Bam*HI-*Sma*I fragments, Fig. 1*A*) were substituted for their wild-type counterparts in plasmid SP64/RchrβG[13], which allows *in vitro* synthesis of a globin pre-mRNA containing both the small and the large intron (Fig. 1*A*). Wild-type, as well as the three mutant pre-mRNAs 1,036 A→G, 1,036 A→C and 1,036 A→T, yielded the correctly-spliced product of 536-nucleotide length when spliced *in vitro* (Fig. 1*B*). The rate of splicing of the large introns of the mutants, as determined within the linear period of the reaction (60-120 min), was about 50% that of wild type (Table 1). In the case of the mutant substrates 1,036 A→G and A→T a slight accumulation of the lariat intermediate of the large intron relative to the lariat product was observed (Fig. 1*B*).

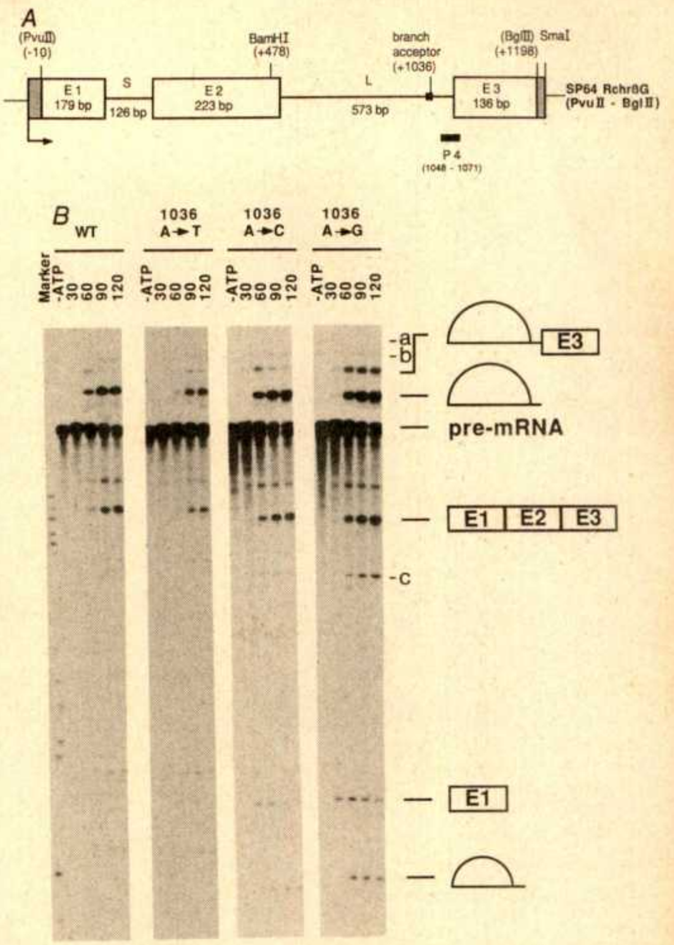


Fig. 1 Analysis of the *in vitro* splicing products of pre-mRNAs with branch acceptor mutations. *A*, Partial map of the *Pvu*II-*Bgl*IIRchrβG DNA fragment[14] cloned into pSP64 for synthesis of pre-mRNA *in vitro*. The construction of the wild-type plasmid has been described[13]. Numbering is relative to the natural cap site (+1). Initiation of transcription is indicated by an arrow. Point mutations at position 1,036 were obtained by oligonucleotide-directed mutagenesis[19]. The mutated *Bam*HI-*Sma*I fragments were substituted for their wild-type counterparts and the resulting plasmids used as template for *in vitro* synthesis of pre-mRNA[22,23]. The resulting transcripts contained 24 and 3 nucleotides of polylinker sequences at the 5′ and 3′ ends, respectively. The location of the oligonucleotide primer used in primer elongation is shown. The drawing is not to scale. *B*, *In vitro* splicing of wild-type and branch-point-mutant pre-mRNA. The ^{32}P-UMP-labelled pre-mRNA (specific activity 2×10^6 c.p.m. pmol^{-1} RNA) of the wild type and the mutants indicated above each panel were incubated with HeLa-cell nuclear extract[6,24] for the time (min.) indicated, and analysed on a 3.2% denaturing polyacrylamide gel[25]. The relevant products and intermediates are indicated; for a description of structures a (superlariat intermediate), b (superlariat) and c (E1-E2), see refs 13 and 19. '-ATP', ATP and creatine phosphate omitted and the RNA incubated with HeLa-cell nuclear extract for 120 min. Marker, 5′-^{32}P-labelled *Bsp*I-digested pBR322.

As shown previously[13], the exact position of the branch point can be determined by primer extension analysis using the 24-nucleotide *primer 4* complementary to the region 1,048-1,071 (ref. 14), which is 12 nucleotides downstream of the branch acceptor A nucleotide (Fig. 1*A*). Reverse transcription of branched RNA molecules, either the excised lariat intron or the lariat intermediate containing the third exon, is arrested at the position of the branch[2]. Wild-type RNA yields a major primer extension product, indicative for the branch at position 1,036 (ref. 13) and a minor product extending to position 1,035, the

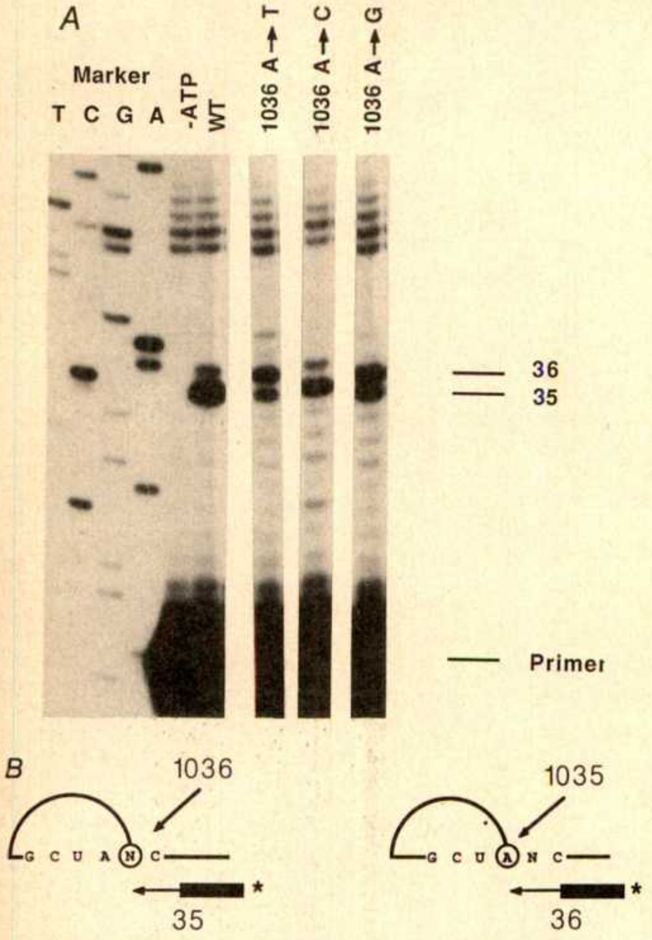


Fig. 2 Analysis of branch point use by primer extension. *A*, Wild-type and mutant pre-mRNA (0.1 pmol each) were incubated with HeLa-cell nuclear extract and RNA was extracted after 1.5 h. For annealing of the 5′-^{32}P-labelled 24-nucleotide *primer 4* (specific activity 4×10^6 c.p.m. pmol^{-1}) and primer extension using reverse transcriptase, see ref. 13. Extension products were analysed on a 12% denaturing polyacrylamide gel[25]. Marker, dideoxy chain termination sequence[26] of single-stranded M13 mp8 RchrβG (*Bam*HI-*Eco*RI) using the universal N15 primer[27]. *B*, Scheme of the expected primer extension products indicative for branch-point formation at position 1,036 and 1,035. Black box, 5′-^{32}P-labelled primer (*, position of label).

Table 1 Relative rate of splicing and use of different branch points

Nucleotide at position 1,036 (N)	Relative rate of splicing	Relative use of branch point 1,036 GCUAⓃ C–	1,035 GCUⒶ NC–
A(wt)	100%	90%	10%
T	60%	35%	65%
C	45%	85%	15%
G	55%	40%	60%

The nucleotide N (in DNA) was substituted with the nucleotides listed in column 1. Appropriate bands were cut out of a polyacrylamide gel like the one in Fig. 1*B* and counted in a liquid scintillation counter. The relative splicing rate was calculated from the ratio of fully spliced product E1-E2-E3 to small lariat intron at different times. The rate of excision of the first intron is independent of the rate of excision of the second intron, as determined by quantifying the data in Fig. 3 of ref. 19, and can therefore serve as a reference value. This ratio is taken as 100% for the wild type. To calculate relative branch point use, the two primer extension products (Fig. 2*A*) indicative for the two branch points were excised from the gel and the radioactivity was determined. wt, wild-type nucleotide.

A residue immediately 5′ to the major branch point (Fig. 2*A*, *B*), in a 10:1 ratio (Table 1). In the case of the mutant pre-mRNAs 1,036 A→T and A→G, about 60% of the branches were displaced to the A residue at position 1,035, while the remaining 40% were connected to the mutated nucleotide at position 1,036 (Fig. 2*A*, Table 1). The 1,036 A→C mutant yielded predominantly C-branched molecules (85%), while the neighbouring upstream A-nucleotide was used in 15% of the lariat molecules. We conclude that all four nucleotides can serve as branch acceptors, albeit with different efficiencies, namely A > C > G, U.

Since in the above experiment a mixture of lariat intermediate and lariat product had been analysed, it was not clear whether the molecules that were branched at the mutated nucleotide were capable of undergoing cleavage and exon joining at the 3′ splice site, or whether they accumulated as dead end lariat intermediates. We isolated the lariat product as well as the lariat intermediate by polyacrylamide gel electrophoresis. In the case of wild type, lariat intermediate and lariat product yielded the same pattern of primer extension products (Fig. 3): the A residue at position 1,036 was the major branch acceptor in both cases. Different results were obtained with the two mutants 1,036 A→G and 1,036 A→T: whereas the lariat products were branched almost exclusively at the A residue (position 1,035) immediately upstream of the mutated position, the lariat intermediate was branched predominantly at the mutated position itself. In contrast, the lariat intermediate and lariat intron derived from the mutant pre-mRNA 1,036 A→C yielded mainly lariat product branched at the mutated residue (Table 2). This analysis assumes that the lariat products of all the pre-mRNAs examined are equally stable in the *in vitro* system and are therefore representative of the processed lariat intermediates. We believe this to be the case because the ratio of lariat product to mature mRNA was about the same for the wild type and all the mutants throughout the reaction (Fig. 1*B*). We have also determined that the deproteinized mutant lariat products are more resistant to debranching by nuclear extract than the wild-type lariat (the relative rates decrease in the order wild type >1,036 A→T, 1,036 A→G > 1,036 A→C), however, as native lariats are resistant to debranching[15], this finding is not necessarily relevant to the question of selective degradation.

We conclude from these results that lariat intermediate branched at a C-nucleotide is functional in the second step of the splicing reaction, whereas G- or U-branches are strongly inhibitory.

This observation is in agreement with the results of Wallace and Edmonds[16], who reported that the predominant branched nucleotide in nuclear RNA was an A residue. In contrast to previous reports, more detailed analysis revealed that branched C-residues were the next most frequently found (J. C. Wallace and M. Edmonds, personal communication). Moreover, a naturally-occurring C-branched lariat of the first intron of the human growth hormone pre-mRNA has been identified (K. Hartmuth and A. Barta, personal communication).

An A→G transition of the branch acceptor of the small intron of the human β-globin gene has been shown to lead to a fivefold reduction of the splicing efficiency *in vitro*[12]. The mutated branch acceptor nucleotide (37 nucleotides upstream of the 3′ splice site) was not used in lariat formation. Instead, a cryptic branch acceptor sequence 24 nucleotides upstream of the 3′ splice site was activated. The difference between these results and ours may be due to the fact that different introns were examined in the two studies. The affinity of the original branch acceptor sequence for the human β-globin small intron to the component recognizing it may be weaker than that of the rabbit sequence we have studied. Further weakening of this affinity by an A→G transition of the branchpoint might lead to the use of a cryptic site which has a more favourable sequence and/or is closer to the 3′ splice site. There seems to be no stringent sequence

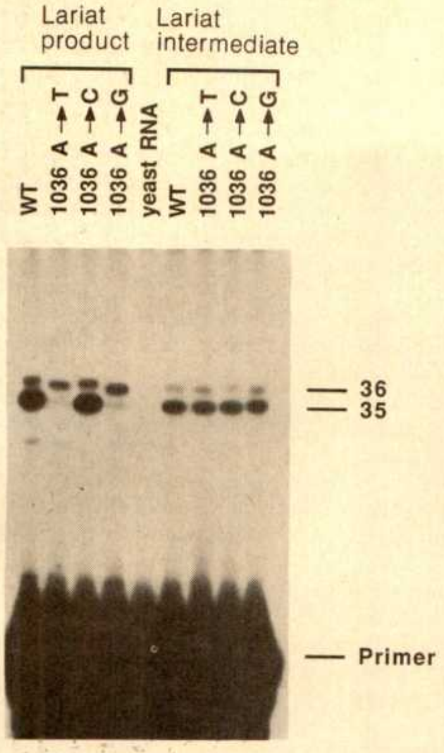


Fig. 3 Branching in lariat product and lariat intermediate derived from pre-mRNAs with branch point mutations, analysed by primer extension. Wild-type and mutant pre-mRNAs (4 pmol each) were spliced *in vitro* as described in Fig. 1. Lariat product and lariat intermediate were isolated by preparative polyacrylamide gel electrophoresis and analysed by primer extension using the 5'-^{32}P-labelled *primer 4* (see Fig. 1*A*). The 35- and 36-nucleotide extension products indicate branch points at positions 1,036 and 1,035, respectively (see Fig. 2*B*). Yeast RNA was used as carrier RNA and served as a negative control.

requirement for branch acceptor sequences, as indicated by the variety of natural and cryptic branch acceptor sequences reported to date[12,13].

All mutations of the branch acceptor A of the rabbit β-globin large intron led to a 50% reduction in the rate of splicing as compared to wild type. Since complex formation around the branch acceptor sequence is an early step in the splicing reaction[17,18], the observed rate in the three mutant RNA splicing reactions may be caused by a lower affinity of the splicing components for the branch acceptor sequence, or by a decreased rate of branch formation. Our results do not allow us to differentiate between these two possibilities. As the branch is nevertheless formed at the mutated position or at the A-nucleotide just upstream of the normal branch site, and no activation of cryptic branch acceptor sequences is detected, we conclude that the mutated sequence is still functional, which confirms that there is no stringent sequence requirement for branch formation. Although an A- or C-residue is favoured over G or U, branch formation is possible at all four nucleotides. In the case of the mutants, preference for an A residue may cause the activation of the A-nucleotide at position 1,035, one nucleotide upstream of the normal branch site, as branch acceptor.

It is surprising that the sequence requirements for the first step of the splicing reaction, where the branch region is intimately involved, are lower than for the second step, where it is at a distance from the reacting nucleotides. In the first step, recognition of the branch-forming region by splicing components may not be stringent, and /or the reactivity of the branch-forming nucleotide as regards its capacity of attacking the phosphodiester bond upstream of the 5' terminal G residue of the intron may not be very sequence specific. It is even possible that assembly of the nucleoprotein complex at the branch-forming region takes place at a fixed distance from the polypyrimidine stretch at the end of the intron, regardless of the actual sequence at that position, and that branch formation occurs from the closest suitable nucleotide.

Table 2 Different frequencies of branchpoints in lariat products and lariat intermediates

Nucleotide at position 1,036 (N)	Relative use of branch point			
	Lariat product		Lariat intermediate	
	1,036 GCUAⓃC	1,035 GCUⒶNC	1,036 GCUAⓃC	1,035 GCUⒶNC
A(wt)	80%	20%	80%	20%
T	10%	90%	80%	20%
C	80%	20%	90%	10%
G	20%	80%	80%	20%

The primer extension analysis of Fig. 3 was quantified by cutting the cDNA corresponding to different branch points out of the gel, and determining the radioactivity in a scintillation counter. wt, wild-type nucleotide.

As regards the second step of splicing, it is not immediately clear why 3' cleavage and exon joining should be affected by a mutation at the branch point, which is some 20-30 nucleotides upstream of the 3' cleavage site. It is important to note that certain mutations at the 5' end of the β-globin large intron, namely those converting the 5' GT to AT, GA and GG, also prevent the second step of splicing, while allowing the first[19]. A similar observation has been made for a yeast 5' GT→AT mutant[20]. In yeast, moreover, the requirements at the branch point (TACTAAC) are apparently even stricter than in higher eukaryotes, because an A→C substitution of the branch acceptor nucleotide not only slows down branch formation to the C-residue but very strongly reduces the second step of splicing[21]. It would thus seem that at least three nucleotides close to the branch, namely the branch point nucleotide and the two closest residues linked to it via the 2'-5' linkage, are required for the second splicing step. This requirement could come about either because a component that binds to the branch region in connection with the first splicing step does not undergo an essential conformational change, or fails to dissociate or translocate or because a component required for the second step of the splicing reaction is unable to recognize the mutated sequence.

Supported by grants from the Schweizerische Nationalfonds and the Kanton of Zürich. H. H. was a recipient of an EMBO Long-Term Fellowship. We thank D. Schuppli for technical assistance and F. Ochsenbein for preparing the photographs.

Received 27 June; accepted 14 October 1986.

1. Padgett, R. A., Konarska, M. M., Grabowski, P. J., Hardy, S. F. & Sharp, P. A. *Science* **225**, 898-903 (1984).
2. Ruskin, B., Krainer, A. R., Maniatis, T. & Green, M. R. *Cell* **38**, 317-331 (1984).
3. Zeitlin, S. & Efstratiadis, A. *Cell* **39**, 589-602 (1984).
4. Domdey, H. *et al. Cell* **39**, 611-621 (1984).
5. Rodriguez, J. R., Pikielny, C. W. & Rosbash, M. *Cell* **39**, 603-610 (1984).
6. Konarska, M. M., Grabowski, P. J., Padgett, R. A. & Sharp, P. A. *Nature* **313**, 552-557 (1985).
7. Langford, C. J. & Gallwitz, D. *Cell* **33**, 519-527 (1983).
8. Pikielny, C. W., Teem, J. L. & Rosbash, M. *Cell* **34**, 395-403 (1983).
9. Langford, C. J., Klinz, F.-J., Donath, C. & Gallwitz, D. *Cell* **36**, 645-653 (1984).
10. Keller, E. B. & Noon, W. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7417-7420 (1984).
11. Reed, R. & Maniatis, T. *Cell* **41**, 119-126 (1985).
12. Ruskin, B., Greene, J. M. & Green, M. R. *Cell* **41**, 833-844 (1985).
13. Padgett, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 8349-8353 (1985).
14. Van Ooyen, A., van den Berg, J., Mantei, N. & Weissmann, C. *Science* **206**, 337-344 (1979).
15. Ruskin, B. & Green, M. R. *Science* **229**, 135-140 (1985).
16. Wallace, J. C. & Edmonds, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 950-954 (1983).
17. Black, D. L., Chabot, B. & Steitz, J. A. *Cell* **42**, 737-750 (1985).
18. Ruskin, B. & Green, M. R. *Cell* **43**, 131-142 (1985).
19. Aebi, M., Hornig, H., Padgett, R. A., Reiser, J. & Weissmann, C. *Cell* (in the press).
20. Newman, A. J., Lin, R. J., Cheng, S. C. & Abelson, J. *Cell* **42**, 335-344 (1985).
21. Jacquier, A. & Rosbash, A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5835-5839 (1986).
22. Melton, D. A. *et al. Nucleic Acids Res.* **9**, 7035-7056 (1984).
23. Konarska, M. M., Padgett, R. A. & Sharp, P. A. *Cell* **38**, 731-736 (1984).
24. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
25. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
26. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
27. Heidecker, G., Messing, J. & Gronenborn, B. *Gene* **10**, 69-73 (1980).

The correct activation of *Antennapedia* and bithorax complex genes requires the *fushi tarazu* gene

P. W. Ingham* & A. Martinez-Arias

Medical Research Council, Laboratory of Molecular Biology,
Hills Road, Cambridge OX1 3PS, UK

In the *Drosophila* embryo the establishment and specification of metameric units depends upon the selective activation of the segmentation[1,2] and the homoeotic selector genes[3,4]. The former are necessary for establishing the appropriate number of metameric[2] or parasegmental[5] units, whereas the latter control the pathways of differentiation followed by particular parasegments[3,4]. Classical embryological manipulations[6,7] have shown that these processes must be closely coordinated during normal development. However, previous studies of pair-rule genes have led to the suggestion that the specification of segmental identity proceeds independently of the establishment of metameres as physical units[1,8]. These apparently conflicting perspectives can be reconciled by envisaging a common maternally derived positional information system which is independently interpreted by the components of both processes[6,9]. In the case of the partitioning process, the gap and pair-rule genes would be instrumental in translating this information, whereas the activation of the homeotic genes would be mediated via other intermediaries[4,10] (see ref. 9 for review). It is difficult to see, however, how such a system could ensure the precise regulation of the two types of genes implicit in the final differentiated pattern. This difficulty has led to the suggestion that the segmentation mechanism must define the precise boundaries of selector gene expression[11]. Here we confirm this suggestion and propose that the gene *fushi tarazu* plays a key role in this process, integrating the processes of metameric partitioning and regional specification in the *Drosophila* embryo.

The gene *fushi tarazu* (*ftz*)[2,12], a member of the pair-rule class, differs from the other members of the class in two interesting ways: it is located in the Antennapedia-complex (ANT-C)[13], one of two major clusters of selector genes, and it contains a class I homoeobox[14,15], which is present in and highly conserved between the ANT-C and bithorax-complex (BX-C) selector genes[14-16]. These characteristics have led to the suggestion that *ftz* is itself a homoeotic gene[17,18] which, in some unspecified way, regulates the expression of the ANT-C and BX-C genes[15]. Here we test these proposals and argue that *ftz* plays a dual role in the *Drosophila* blastoderm; in common with other pair-rule genes, it is necessary for the correct spatial regulation of segment-defining genes[19] for example *engrailed*, (*en*). In addition, we present evidence that it is necessary for the correct activation of homoeotic genes.

During normal embryogenesis the establishment of parasegmental units is first seen by the appearance of fourteen stripes of *en* activity along the antero-posterior axis of the early gastrula[19-21]. Shortly afterwards these stripes come to lie at the anterior margin of fourteen metameric units which can thus be defined as parasegments (PS)[5,22] (Fig. 1*a*). The specification of parasegments is illustrated at this stage by the distribution of selector gene products[5,16,23-26,28,29]. The *Ubx* proteins, for example, show a characteristic heterogeneous pattern between PS5 and PS13[27], defining a metameric repeat or 'Ubx-metamere'[28]; in the wild type these metameres coincide with parasegments and PS6 displays a characteristic higher intensity and more even distribution of proteins than other parasegments (see Fig. 1*b*).

In the extended germ band, *ftz*$^-$ embryos are partitioned into seven double-width parasegments[17] and lack all the even-numbered engrailed stripes[19] (Fig. 1*c*). At the same stage, embryos mutant for another pair-rule gene *odd paired* (*opa*), are similarly malpartitioned, though vestiges of the even numbered stripes of engrailed do remain (S. DiNardo and P. O'Farrell, personal communication; Fig. 1*e*). If the specification of parasegments is independent of their establishment, as previously suggested[1,8,9], the malpartitioning observed in *ftz*$^-$ and *opa*$^-$ embryos would be expected to generate parasegments of compound identity (see Fig. 2). To investigate this possibility we have analysed the distribution of *Ubx* proteins in *opa*$^-$ and *ftz*$^-$ embryos.

In embryos mutant for *opa*, the pattern of *Ubx* protein distribution in the double-sized units reveals them to be of composite character, the anterior half behaving as an odd-numbered parasegment and the posterior as an even-numbered parasegment. This is particularly clear in the case of the PS5/6 unit, in which the posterior half exhibits high levels of *Ubx* protein typical of the wild type PS6 (Fig. 1*f*). Each of the three more posterior double units displays defects in the *Ubx* pattern that are attributable to *opa* effects on segment polarity genes (unpublished data).

By contrast, the *Ubx* protein pattern in *ftz*$^-$ germ bands is quite different both from *opa* and wild type. The PS5/6 double unit has a staining pattern typical of the wild type PS5, except for a narrow strip of cells at the posterior border that show levels characteristic of a more posterior parasegment (but not PS6). The next double unit exhibits *Ubx* staining characteristic not of PS6 but of PS7-12, as so the next two double units, whereas the last double unit resembles PS13. The main feature of this pattern is that there are half the number of parasegments and half the number of Ubx-metameres. Both units coincide, and the *ftz*$^-$ embryos can be considered to consist of a chain of large odd-numbered parasegments: 3*, 5*, 7*, 9*, 11*, 13* (Fig. 1*d*).

Thus the two mutations *opa* and *ftz*, which result in similar malpartitioning of the germ band, show radically different patterns of *Ubx* expression within three hours of the blastoderm stage. The absence of PS6-specific *Ubx* staining in *ftz* germ bands cannot be accounted for by the elimination of PS6 cells, as little if any cell death has occurred in the ectoderm by this stage (unpublished observations). Rather, it suggests a failure in the specification of the metameric unit at blastoderm. We therefore monitored the initiation of *Ubx* expression in *ftz*$^-$ blastoderms.

An interesting and unexpected feature of the initial expression of *Ubx* in normal embryos is a transient pair-rule distribution of *Ubx* transcripts reminiscent of and in the same frame as that of *ftz*[11,23] (see Fig. 3). This discontinuous pattern is superimposed on a broader continuous distribution of transcripts. *Ubx* transcripts do accumulate in the *ftz*$^-$ blastoderm, but show a continuous distribution, with some modulation in the dorsal region (Fig. 3), but without the peak of accumulation ventrally or laterally in the anlage of PS6 that is the first sign of the pair-rule modulation in the wild type. Accumulation of transcript begins slightly posterior to the position of PS6 and probably reflects the wild-type early continuous distribution. At onset of gastrulation in wild-type embryos three further peaks of *Ubx* transcript[11,23] appear which correspond to PS8, PS10 and PS12 (Fig. 3); in *ftz*$^-$ embryos *Ubx* expression remains homogeneous at this stage. By the end of gastrulation, when the pair-rule modulation is normally still clearly visible, the pattern of *Ubx* expression reveals three unusually large and similar units reflecting the malpartitioning of the embryo and probably corresponding to the 7*, 9* and 11* parasegments described above (Fig. 3*f*).

To determine whether there is a requirement for *ftz*$^+$ in the activation of other selector genes we have studied the initiation of expression of the *Sex Combs Reduced* (*Scr*) and *Antennapedia* (*Antp*) loci in *ftz*$^-$ embryos. Expression of both genes also shows a transient pair-rule modulation in the late blastoderm[25,26]. In the wild type there is a strong peak of *Scr* expression in PS2

---

* Present address: Imperial Cancer Research Fund, Developmental Biology Unit, Department of Zoology, South Parks Road, Oxford, OX1 3PS, UK.

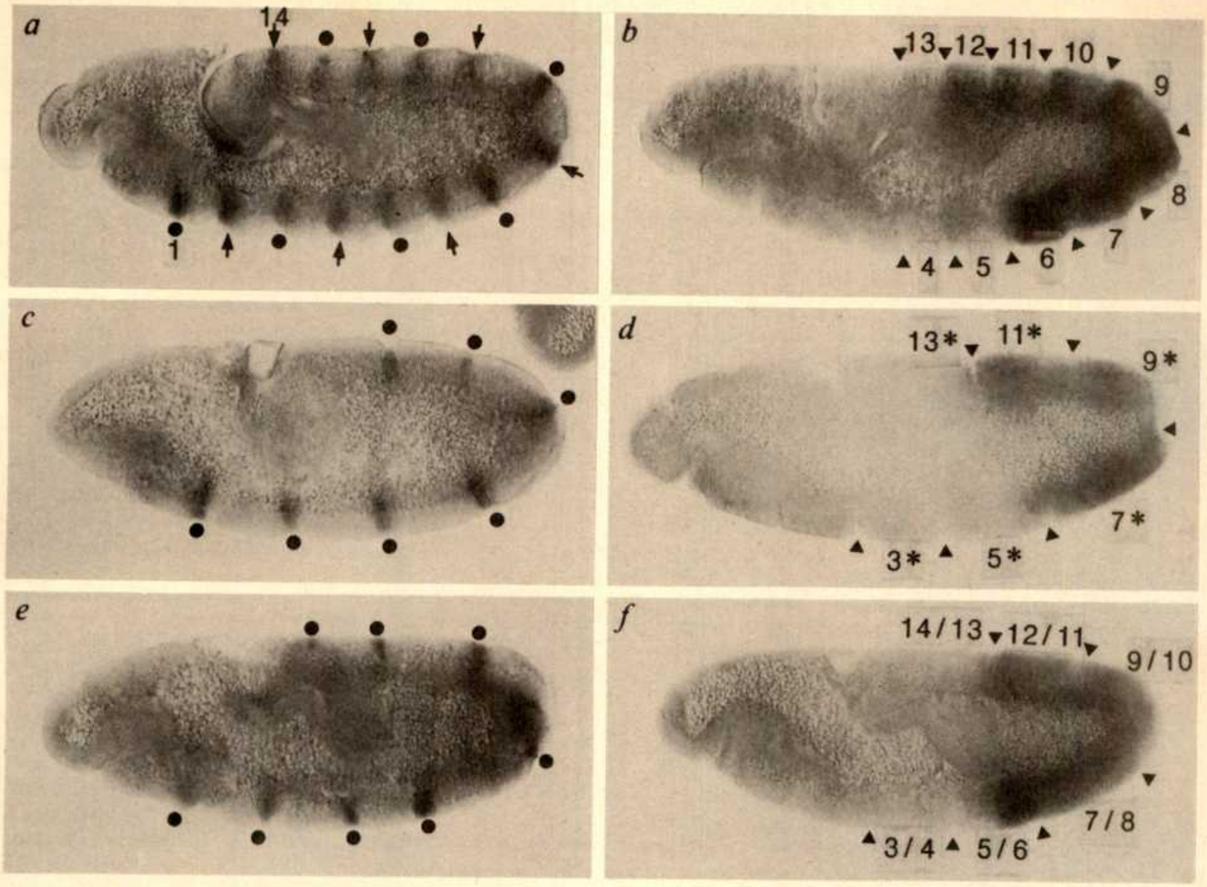


Fig. 1 Coupled and uncoupled partitioning and specification of parasegments visualized by *en* and *Ubx* protein distribution in wild-type and mutant embryos. *a*, Wild-type extended germ band showing bands of *en* protein (1–14), each of which is located at the anterior margin of the odd numbered (small dots) and even numbered (arrows) parasegmental units; *b*, wild type extended germ band showing typical *Ubx* protein distribution. The highest levels of staining are in PS6; PS7–12 show very similar patterns with higher levels posteriorly than anteriorly. PS5 has only limited expression around the tracheal pit and expression in PS13 is also restricted[23,27]; *c*, ftz^- extended germ band stained with *en* antibody. The embryo is subdivided into seven double sized units and only the odd-numbered *en* stripes are present; these are of approximately the same width as wild type, in agreement with previous transcriptional data[19]. *d*, embryo similar to that shown in *c* stained with antibody for *Ubx*. There is no staining characteristic of the wild type PS6. The most anterior unit which expresses *Ubx* appears as an enlarged PS5, with the exception of a few cells at its posterior margin which show a higher level of staining. The following three units have patterns typical of single parasegments, which would correspond to 7*, 9* and 11*; *e*, opa^- extended germ band stained with *en* antibody. Like the ftz^- germ band, it is composed of half the normal number of double width parasegmental units, only of the odd-numbered *en* bands are present. However, vestiges of even-numbered bands are often visible, especially in the mesoderm. *f*, Embryo similar to that shown in *e*. A single segment periodic pattern of *Ubx* staining similar to that of wild type is superimposed on the malpartitioned embryo which is composed of half the wild-type number of double segment units. Note especially that the PS5–6 unit shows staining characteristic of the wild type PS5 and PS6. The deviations from wild type in the more posterior parasegments can be explained by the effects of segment polarity genes on *Ubx* (unpublished observations).

Methods. Embryos were stained with monoclonal antibodies against *en*[20] and *Ubx*[41]. The staining procedure is similar to that described elsewhere[27], except that the second antibody was biotinylated, allowing the use of peroxidase conjugated to a biotin-avidin system (Vectastain kit; Vector laboratories). The stain was visualized with diamino benzidine according to instructions from the suppliers. In these and subsequent experiments, we used three alleles of *ftz*: a deficiency *Df(3R)4Scb*, a point mutant, ftz^{9H34} (ref. 42), and an insertion ftz^{W20} (ref. 18). In the case of *odd paired* a single allele opa^{5H97} (ref. 42) was used.

and six weaker peaks corresponding to even-numbered parasegments. This initial pattern of *Scr* expression is completely abolished in ftz^- embryos (Fig. 4*c*).

Expression of *Antp* from the P2 promoter is first detectable in a four-cell-wide peak which corresponds in position to the primordium of PS4 in the late blastoderm[26], this peak is not present in ftz^- embryos at the same stage (Fig. 4*b*). Transcripts from the P2 promoter only appear in ftz^- embryos during late gastrulation, when they are restricted to the presumptive ectoderm in a region corresponding to PS3 and PS4 (data not shown). This is similar to *Antp* expression in wild-type PS3[26] and is likely to correspond to PS3*. The pattern of *Antp* expression from the P1 promoter is not altered in ftz^- embryos (unpublished observations).

These results suggest that *ftz* plays a fundamental role in differentiating between the even- and odd-numbered parasegments by modulating the activation of the selector genes in the blastoderm. Although the unique identity of each parasegment is controlled by selector genes, the genes themselves are generally expressed in more than one parasegment[16,23–29]. A striking feature of the presumptive gnathal and thoracic region of the embryo at blastoderm (and the central nervous system during later embryonic stages), is that the even-numbered parasegments are the primary sites of expression of different selector genes[23–26]. Thus *Scr* is expressed at highest levels in PS2, *Antp* in PS4 and *Ubx* in PS6; mutations in each of these genes transform their preferred parasegments but not the other even-numbered ones. By contrast, both genetic analysis and expression patterns suggest that the odd-numbered parasegments depend on more than one selector gene for their correct development. Different parasegments thus come to be distinguished both by different combinations and different levels of selector gene activity. Both PS4 and PS5, for instance, express *Antp*, but this expression differs between the two, both in spatial pattern and amount. This

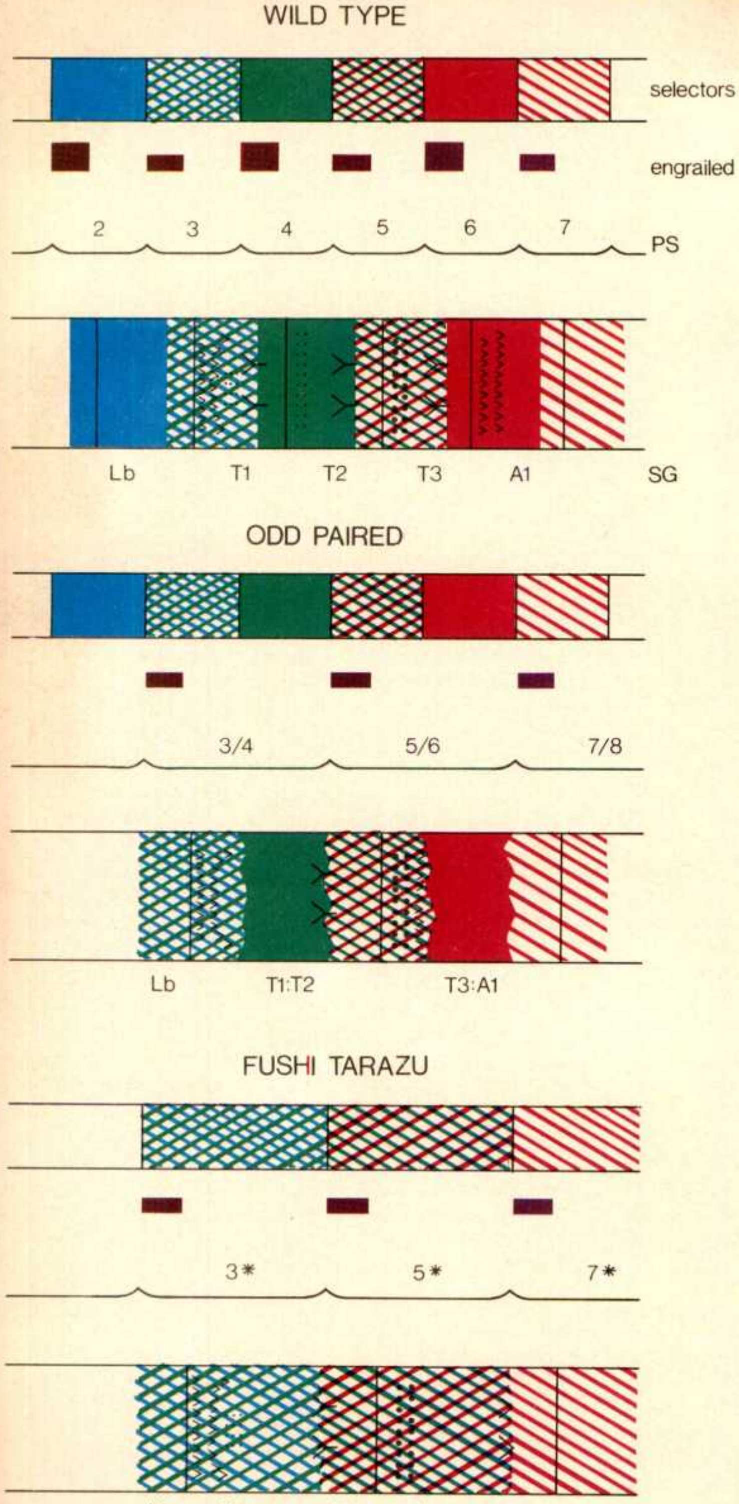


Fig. 2 Schematic representations of the embryonic development in wild type and each of the mutants. Colours indicate activity of selector genes: blue, *Scr*; green, *Antp* and red, *Ubx*. Cells in even-numbered parasegments express a single selector gene in a rather homogeneous pattern. Odd numbered parasegments display complicated patterns with more than one selector gene expressed at high levels in areas not demarcated by compartment boundaries[23-26] (indicated by hatching). Notice how segments are derived from parasegments by shifting the position of the grooves; this occurs during muscle attachment after germ band shortening[5]. In the wild type, segments are thus composed of two adjacent parasegments. In the mutants, this is also true but the selector gene distribution in them results in enlarged segments with rather different identities from wild type. The two different mutant situations illustrated will result in different cuticular phenotypes. In the case of *opa*, where partition and specification are uncoupled, anterior compartments are composed of cells of two different identities; this in turn generates larval cuticular denticle bands of composite segmental character, as observed in *opa*$^-$ larvae[2]. The wavy boundaries represent the observed variability of the phenotype. In the case of *ftz*, where partition and specification are coupled, anterior compartments have a single identity; the final phenotype thus corresponds to the precise elimination of alternate parasegments[2]. Notice that our observations and interpretations contrast sharply with former interpretations of these phenotypes[2], which suggest that the phenotypic differences between *opa* and *ftz* directly reflect a requirement in different but overlapping regions; this requirement would correlate with the site of expression of the gene. In our view, the phenotypes reflect the fundamental differences in the functions rather than the spatial depolyment of both genes.

difference is maintained by *Ubx*, as in the absence of *Ubx*, PS5 (and PS6) now develop like PS4 and express *Antp* accordingly[30,31]. These differences between *Antp* expression in the primordia of PS4 and PS5 are, however, apparent in the late blastoderm stage, long before *Ubx* product could possibly act and indeed in the absence of the *Ubx* gene[29]. It follows that the parasegmental modulation of *Antp* must be established independently of *Ubx* and only subsequently maintained by interactions between selector genes. Similarly *Scr*, which is required and expressed in PS2 and PS3, is expressed at high levels only in the PS2 primordium at blastoderm, whereas *Ubx* is expressed at high levels in the PS6 primordium at the same time.

The peaks of selector gene activity thus seem to differentiate between the even-numbered and odd-numbered parasegmental domains from the onset of their expression. The results presented here show that the patterns of *Scr*, *Antp* and *Ubx* all depend on *ftz* activity; in the absence of *ftz* products none of the peaks are established. This does not prevent the subsequent expression of any of these genes, but the absence of the initial modulation is reflected in the later failure to establish differences between PS3 and 4, 5 and 6 and so on in *ftz*$^-$ embryos. A prediction of this interpretation is that expression of *Scr* and *Antp* extends further posteriorly than normal, so that *Scr* now becomes expressed in cells which would normally have given rise to PS4,

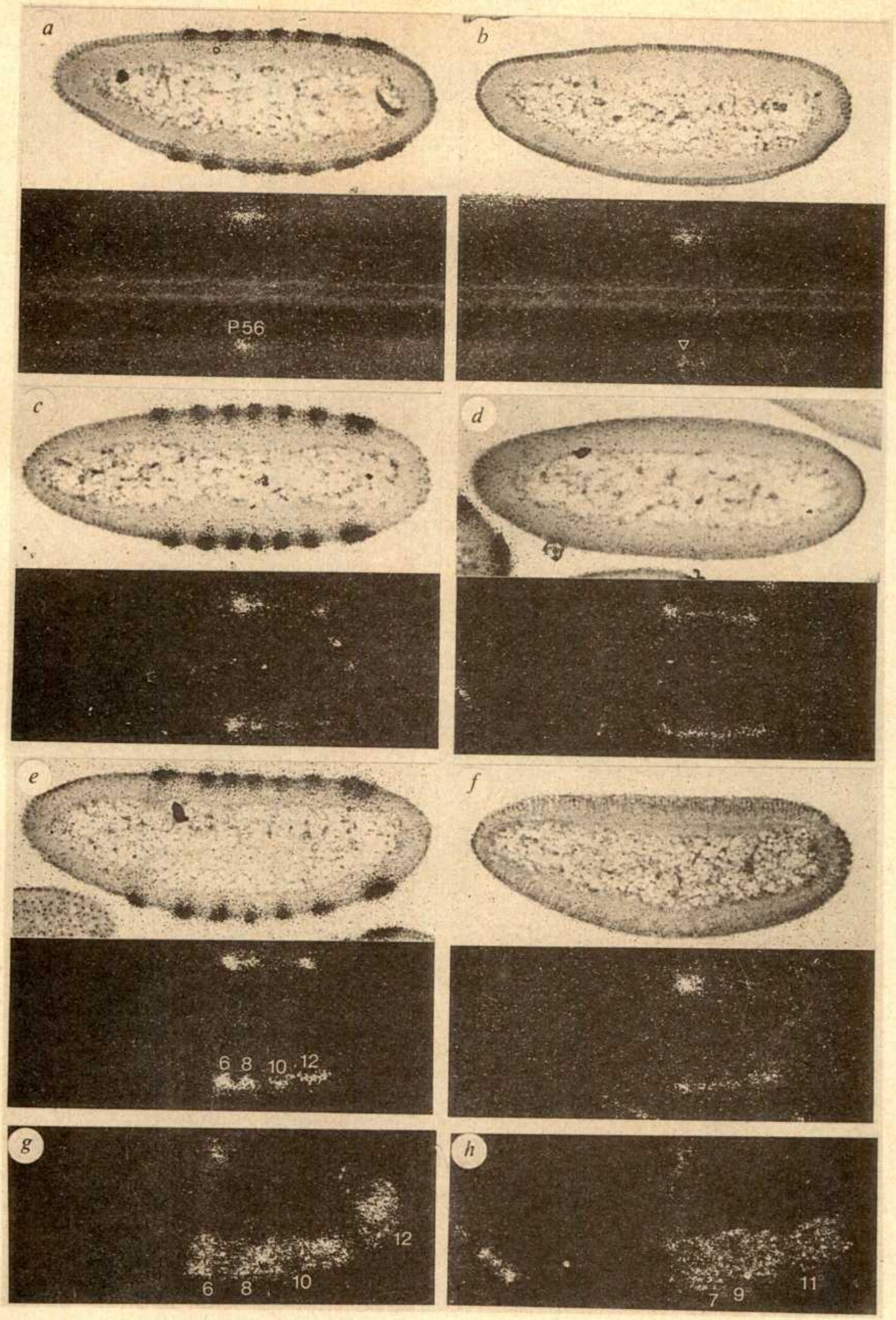


Fig. 3 Initiation of *Ubx* transcription in wild type and *ftz*$^-$ embryos. Panels *a–f* show bright field images of sections hybridized with ^{35}S-labelled *ftz* probe revealing the *ftz* embryos and dark field images of adjacent sections hybridized with ^{35}S-labelled *Ubx* probe. *a*, Wild type; a peak of *Ubx* transcript is first detectable dorsally and ventrally in a region corresponding to PS6; *b*, *ftz*$^-$ embryos at the same stage; there is no clear peak of transcript ventrally and laterally (arrowed). *c*, Slightly older wild type embryos show rather high levels of *Ubx* transcript in PS6, with a low background of transcription extending posteriorly; *d*, no such peak is observed laterally or ventrally in *ftz*$^-$ embryos. *Ubx* transcript is however detected at low levels in a broad domain from 20% to 50% egg length. *e*, With the onset of gastrulation in wild type three further peaks appear[11,23], corresponding to PS8, 10 and 12; *f*, in *ftz*$^-$ embryos, however, the expression ventrally remains more or less uniform. *g*, *h*, dark field images of *Ubx* labelling in late gastrulae. In wild type the pair-rule modulation persists, the label revealing the presumptive parasegments 6, 8, 10 and 12. In *ftz*$^-$ embryos, by contrast, the label reveals three adjacent units which probably correspond to the 7*, 9* and 11* parasegments shown in Fig. 3. The *in situ* methods used are similar to those described elsewhere[43]. The *ftz* probe spans all the *ftz* coding region[43]; the *Ubx* probe corresponds to the A122 fragment[23], and in Northern blots detects a 4.7-kilobase (kb) RNA.

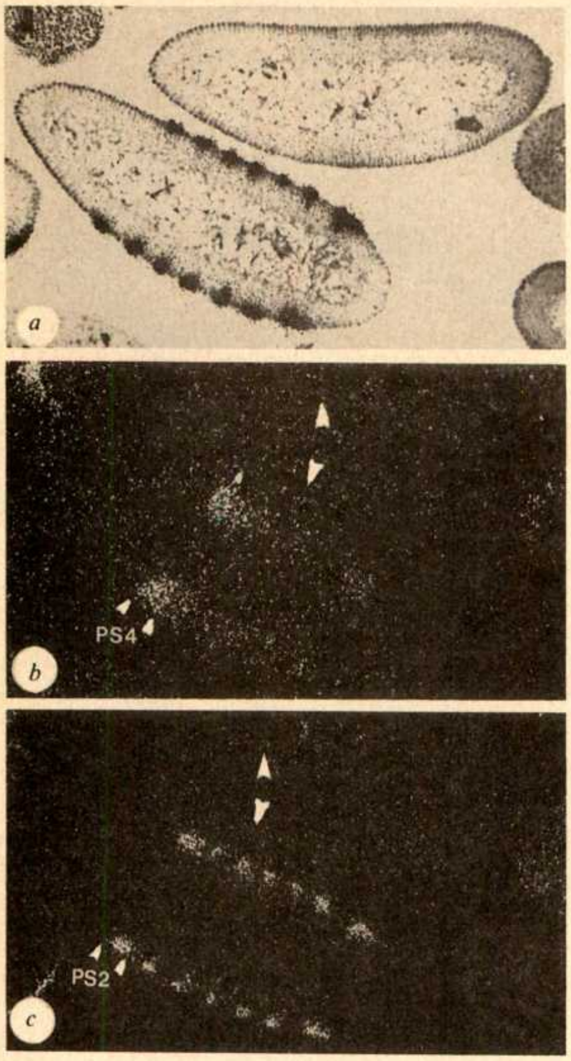


Fig. 4 Expression of *ftz*, *Antp* and *Scr* in wild type and *ftz*^{W20} blastoderms. *a*, Saggital sections through wild type (lower) and *ftz*^{W20} (upper) late stage 14 embryos hybridized with ^{35}S-labelled *ftz* probe. The wild type shows the typical pattern of seven discrete bands of transcript accumulation. The mutant shows very low levels of transcript accumulation which show a weak pair-rule modulation. *b*, Dark field image of adjacent section to *a* hybridized with a ^{35}S-labelled probe specific for the *Antp* P2 transcript. This reveals accumulation of high levels of transcript in the cells of the PS4 primordium in the wild type, but not in the same region of the *ftz*^{W20} embryo (arrowed). *c*, Dark-field image of adjacent section to *a* hybridized with a ^{35}S-labelled probe derived from the 5' end of the *Scr* transcription unit. In the wild type, *Scr* transcript accumulates in seven stripes which correspond in their position and size to those of *ftz* transcript. Transcript accumulation is higher in the most anterior of these, which corresponds to the primordium of PS2. In the *ftz*^{W20} embryo there is no detectable expression of *Scr* at this stage, either in PS2 (arrowed) or more posteriorly. The *Antp* probe is a fragment from the proximal part of the gene; at blastoderm it specifically detects transcripts from the P2 promoter[26]. Methods as in Fig. 3.

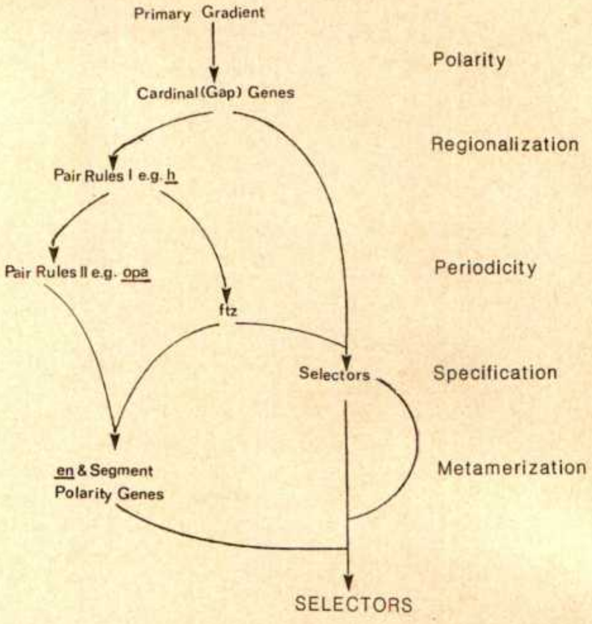


Fig. 5 A summary of the hierarchial interactions which lead to the integrated partitioning and specification of the *Drosophila* embryo. Arrows indicate the regulatory relationships between different classes of genes. The processes proposed to be accomplished by each step are indicated on the right. A primary gradient of maternally derived morphogen(s) establishes 'polarity' and, as the blastoderm begins to form, instructs the deployment of the cardinal or gap genes[40]. This results in the 'regionalization' of the embryo, that is, the establishment of the broad domains of potential selector gene expression. The cardinal genes also act to establish 'periodicity' by organizing the expression of a subset of the pair-rule genes, exemplified by *hairy* (*h*)[32]. The function of this class is to pattern *ftz*, and also, we surmize, other pair-rule genes such as *opa*. This step-by-step subdivision of the embryo is translated into an array of parasegments that are uniquely specified, primarily by the activation of both *en* and the ANT-C and BX-C genes by *ftz*. The 'specification' and 'metamerization' of parasegments takes place as the blastoderm forms into cells. Later, the interactions indicated elaborate the pattern of selectors to that of the first instar larva. This elaborated pattern of selector gene expression is represented by capital letters.

whilst *Antp* is expressed in cells which would normally make PS6. This implies that in the wild type the peaks of expression serve to differentiate between parasegments by establishing specific levels or states of selector gene activity in the cells of the even-numbered parasegments. These states subsequently make these cells refractory to the expression of the more anteriorly expressed selector genes. This close linkage of the selector gene pattern with *ftz* expression is emphasized by the finding that it is maintained in embryos mutant for the gap genes *Kr*, *kni* and *hb* all of which alter the pattern of *ftz* expression[32,33].

Previous analyses have demonstrated a functional non-equivalence of the pair-rule genes, some being required for normal expression of *ftz* whilst others are not[19,34]. These findings led to the proposition that parasegments are established by a hierarchy of interactions between pair-rule genes resulting in the establishment of the pattern of *en* expression[19]. We propose a further subdivision of the pair-rule class distinguishing between those genes that are responsible only for the partitioning process (for example, *opa*) and *ftz*, which acts both to partition and to specify the cells of the blastoderm. Thus, the processes of partitioning and specification of the *Drosophila* embryo are integrated through a hierarchical network of interactions between the different segmentation genes (see Fig. 5).

As *ftz* itself is expressed in a reiterated pattern throughout the presumptive germ band, it is clear that it cannot alone account for the differential activation of the selector genes. There must therefore be specific activators which allow *ftz* to discriminate between the selector genes in different regions of the blastoderm; candidates for such activators are the zygotic gap or cardinal genes[1,2] whose products are distributed in broad though restricted domains[35] and whose absence has been shown to affect selector gene expression[27,32,33]. Consequently, understanding the correlation between the metameric pattern and selector gene expression demands a consideration of the possible molecular interactions between *ftz* and the cardinal gene products.

The data available[32,33] suggest that an interaction between *ftz* and cardinal gene products is responsible for the activation of selector genes, the *ftz* product providing the precise spatial specificity and the gap gene products the promoter specificity. One possibility is that *ftz* and the cardinal products are activators; the specificity of selector gene expression would then result from a simple combination of their activities. An alternative envisions a maternally encoded repressor common to all selector genes and bound at the sites of cardinal gene activity. The role of *ftz* would be to compete out this repressor and thus make the underlying pattern of cardinal gene activity available to selector genes; in this possibility the role of *ftz* in activation, though crucial, would be indirect.

Our model contrasts with those previously proposed in which a primary gradient, maternally encoded, is directly responsible for the activation of the homoeotic selector genes[4,39]. Although there is good evidence for the existence of such gradients[6,36–38], it seems unlikely that they can have the instructive capacity required for the short range detailed patterning that precisely orders selector genes in the zygote. They can, however, as suggested by Meinhardt[40], have instructive capacity in broader, less precisely defined ranges, and may activate the cardinal genes. Thus, the ultimate effect of such gradients on the selector genes would operate through the aforementioned hierarchy (Fig. 5) and ultimately rely on their indirect effect on *ftz*.

We thank M. Akam for encouragement and many critical discussions, C. Cabrera, D. Hogness, K. Howard, P. Lawrence and K. Nasmyth for interest and discussions, C. Nüsslein-Volhard, G. Jürgens and T. Kaufman for mutant stocks and M. Scott for ANT-C probes. We thank M. Eggo, B. Cross and A. Donoughue for typing. A.M.-A. was supported by an EMBO long-term fellowship.

Note added in proof: Results compatible with our conclusions have recently been reported by Duncan, I. M. *Cell*, **47**, 297–309 (1986).

Received 6 October; accepted 17 November 1986.

1. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795–801 (1980).
2. Nüsslein-Volhard, C., Wieschaus, E. & Jürgens, G. *Verh. dt. zool. Ges.* 91–104 (1982).
3. Garcia-Bellido, A. *Ciba Fdn Symp.* **29**, 161–182 (1975).
4. Lewis, E. B. *Nature* **276**, 565–570 (1978).
5. Martinez-Arias, A. & Lawrence, P. A. *Nature* **313**, 639–642 (1985).
6. Sander, K. *Adv. Insect Physiol.* **12**, 125–238 (1976).
7. Newman, J. M. & Schubiger, G. *Devl Biol.* **79**, 128–138.
8. Sander, K., Lohs-Schardin, M. & Baumann, N. *Nature* **287**, 841–843 (1980).
9. Lawrence, P. A. *Cell* **26**, 3–10 (1981).
10. Struhl, G. *Nature* **293**, 36–41 (1981).
11. Akam, M. E. *Phil. Trans. R. Soc.* **B312**, 179–187 (1985).
12. Wakimoto, B. T. & Kaufman, T. C. *Devl Biol.* **81**, 51–64 (1981).
13. Lewis, R. A., Kaufman, T. C., Denell, R. E. & Tallerico, P. *Genetics* **95**, 367–381 (1980).
14. Scott, M. P. & Weiner, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4115–4119 (1984).
15. Laughon, A. & Scott, M. P. *Nature* **310**, 25–31 (1984).
16. McGinnis, W., Levine, M., Hafen, E., Kuroiwa, A. & Gehring, W. J. *Nature* **308**, 428–433 (1984).
17. Wakimoto, B. T., Turner, I. R. & Kaufman, T. C. *Devl Biol.* **102**, 147–172 (1984).
18. Weiner, A. J., Scott, M. P. & Kaufman, T. C. *Cell* **37**, 843–851 (1984).
19. Howard, K. R. & Ingham, P. W. *Cell* **44**, 949–957 (1986).
20. DiNardo, S., Kuner, J. M., Theis, J. & O'Farrell, P. H. *Cell* **43**, 59–69 (1985).
21. Kornberg, T., Siden, I., O'Farrell, P. & Simon, M. *Cell* **40**, 45–53 (1985).
22. Ingham, P. W., Martinez-Arias, A., Lawrence, P. A. & Howard, K. R. *Nature* **317**, 634–636 (1985).
23. Akam, M. E. & Martinez-Arias, A. *EMBO J.* **4**, 1689–1700 (1985).
24. Martinez-Arias, A. *EMBO J.* **5**, 135–141 (1985).
25. Martinez-Arias, A., Ingham, P. W., Scott, M. P. & Akam, M. E. *Development* (submitted).
26. Martinez-Arias, A., Scott, M. P. & Akam, M. E. (in preparation).
27. White, R. A. H. & Lehmann, R. *Cell* **47**, 311–321 (1986).
28. White, R. A. H. & Wilcox, M. *EMBO J.* **4**, 2035–2044 (1985).
29. Harding, K., Weeden, C., McGinnis, W. & Levine, M. *Science* **229**, 1236–1242 (1985).
30. Hafen, E., Levine, M. & Gehring, W. J. *Nature* **307**, 287–289 (1984).
31. Carroll, S., Laymon, R. A., McCutcheon, A., Riley, P. D. & Scott, M. P. *Cell* **47**, 113–122 (1986).
32. Ingham, P. W., Ish-Horowicz, D. & Howard, K. R. *EMBO J.* **5**, 1659–1665 (1986).
33. Martinez-Arias, A., Akam, M. E. & Ingham, P. W. (in preparation).
34. Carroll, S. B. & Scott, M. P. *Cell* **45**, 113–126 (1986).
35. Knipple, D., Seifert, E., Rosenberg, U., Preiss, A. & Jäckle, H. *Nature* **317**, 40–44 (1985).
36. Lehman, R. thesis, Tübingen Univ. (1985).
37. Mlodzik, M., Fjose, A. & Gehring, W. J. *EMBO J.* **4**, 2961–2969 (1985).
38. Schüpbach, T. & Wieschaus, E. *Wilhelm Roux Arch. dev. Biol.* **195**, 302–317 (1986).
39. Capdevila, M. P. & Garcia-Bellido, A. *Wilhelm Roux Arch. dev. Biol.* **190**, 1339–350 (1981).
40. Meinhardt, H. *J. Cell Sci. Suppl.* **4**, 357–387 (1986).
41. White, R. A. H. & Wilcox, M. *Cell* **39**, 163–171 (1984).
42. Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 283–296 (1984).
43. Ingham, P. W., Howard, K. R. & Ish-Horowicz, D. *Nature* **318**, 439–445 (1985).

Does nuclear integration of mitochondrial sequences occur during senescence in *Podospora*?

France Koll

Centre de Génétique Moléculaire, Centre National de la Recherche Scientifique, 91190 Gif-sur-Yvette, France

Observations in various organisms[1–3] suggest that the transfer of mitochondrial DNA sequences to the nucleus has occurred in the course of evolution. Wright and Cummings have reported[4], on the basis of hybridization experiments, that there is such a transposition in the fungus *Podospora anserina* during senescence[5] (arrest of vegetative growth) which is accompanied by amplification of specific mitochondrial sequences[6–9]. They suggested that this transposition could explain senescence, and, as it could be regularly observed under laboratory conditions, senescence in *Podospora* could constitute a model system to study the molecular mechanisms of the transfer of DNA sequences from mitochondria to the nucleus. We report here experiments similar to those of Wright and Cummings[4] but include additional critical controls. Our results do not confirm the data of Wright and Cummings; the only positive signals we observed after hybridization of mitochondrial DNA sequences on nuclear DNAs can be attributed to contaminating DNA from mitochondria or plasmids.

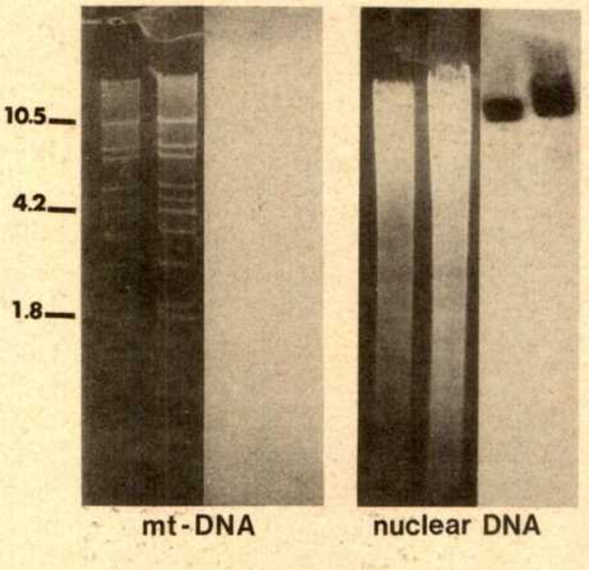


Fig. 1 Detection of a unique nuclear sequence. *a*, Wild-type strain; *b*, *mex1* strain. Nuclear and mt DNAs obtained from a total DNA extract[15] after separation on diamidino-phenyl-indole (DAPI)-CsCl buoyant density gradient[7] (3 or 4 cycles). The 1.4-kb nuclear sequence used as probe contained the suppressor gene *su4* of *Podospora* previously selected from a genomic library[19] and was nick-translated[20] with [α-^{32}P]dATP and [α-^{32}P]dCTP (specific activity $\sim 5\times10^7$ c.p.m. μg^{-1}), hybridized as already described[15] with nitrocellulose filters on which restriction digests[21] of the DNAs were transferred according to Southern[22]. Shown here are the autoradiograms observed after hybridization of the 1.4-kb nuclear sequence on the *Bam*HI digestion of ~5 μg nuclear and ~0.5 μg mt-DNA from a young wild-type strain and from the long-lived mutant *mex1* (nuclear DNA complexity: 2×10^7 base pairs (bp); mt-DNA complexity: 9×10^4 bp). The electrophoretic patterns of the DNAs from the wild-type strain are represented in *a* and from the *mex1* strain in *b*. The autoradiograms (2 days exposure at room temperature) are represented in *a'* and *b'*. We detect an 11-kb sequence similar to the probe in the nuclear DNA of each strain. The probe does not hybridize on the mt-DNA. Sizes in kb.

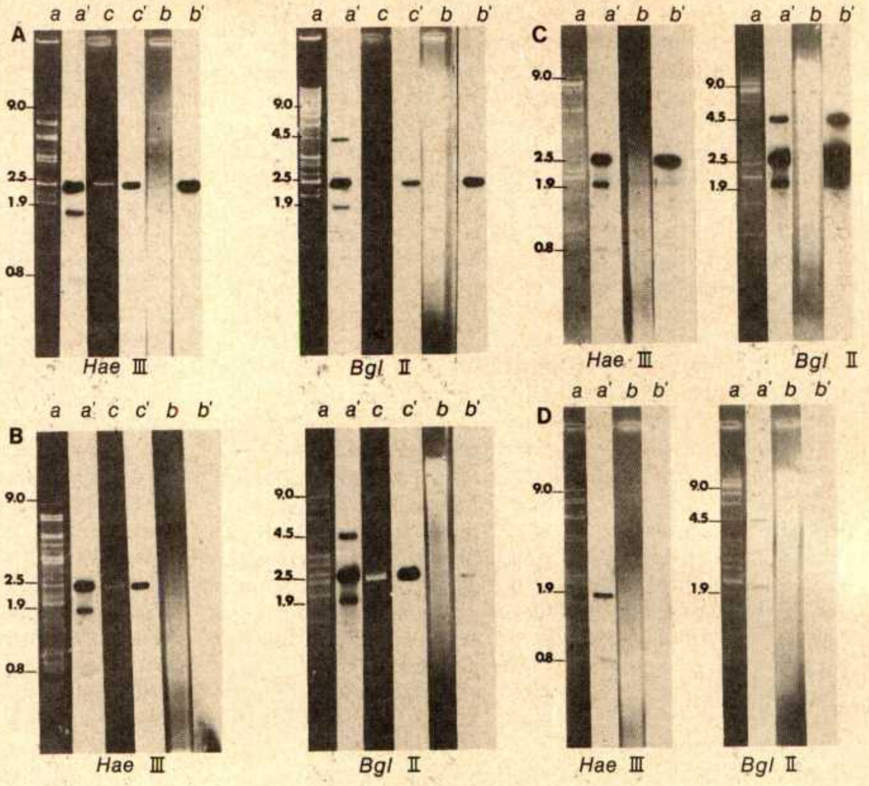


Fig. 2 Hybridization of the α mitochondrial intron on the different DNAs extracted from three independent senescent cultures (A, B, C) and one young *s mt+* (D). Nuclear DNA (density, 1.712 g cm^{-3}), α-free sen-DNA molecules (density, 1.700 g cm^{-3}) and mt DNA (density, 1.694 g cm^{-3}) of each senescent culture were purified by at least three cycles of DAPI–CsCl ultracentrifugation of total DNA extract. Each DNA was digested by *Hae*III and *Bgl*II, electrophoresed on agarose gels, transferred to nitrocellulose filters and hybridized with a ^{32}P nick-translated α purified sequence. The α sequence was purified from the plasmid in which it was integrated by *Bam*HI digestion, electrophoresis on acrylamide gel and electroelution[15]. For both enzymes (*Hae*III, *Bgl*II), the electrophoretic pattern of mt-DNA, α-free sen-DNA molecules and nuclear DNA from a given culture are represented respectively in lanes *a*, *b*, *c* after staining by ethidium bromide (α-sen-DNA extracted from the culture C is not represented). Corresponding autoradiograms after hybridization are shown in lanes *a'*, *b'* and *c'*. For mt-DNA and α-sen-DNA, autoradiograms were exposed for 24 h at room temperature; for nuclear DNAs, they were observed after 4–5 days at −70 °C using two intensifying screens. D, Control experiment showing hybridization of the α sequence on *Hae*III and *Bgl*II digests and corresponding autoradiograms of mt-DNA (*a* and *a'*) and nuclear (*b* and *b'*) DNAs from a young *s mt+* wild type strain. The α mt-sequence used as probe is 2.5-kb long and contains only one site for both enzymes *Hae*III and *Bgl*II. Two fragments (1.9 and 0.8 kb after *Hae*III digest; 4.5 and 1.9 kb after *Bgl*II digest), corresponding to the α sequence integrated in the mitochondrial genome, are labelled after hybridization of the probe on the young mt-DNA (D, lanes *a'* and *b'*). The free α-sen-DNA molecules present in the senescent cultures are circular molecules constituted of tandemly repeated α sequence. Consequently, only one 2.5-kb fragment corresponding to linearized α sequence is observed after digestion of this α-*sen*-DNA by *Hae*III and *Bgl*II (A and B, lanes *c*) and hybridized with the probe (A and B, lanes *c'*). The three labelled bands observed on *Hae*III and *Bgl*II digestion of senescent mt-DNA (parts A, B and C, lanes *a'*) correspond to hybridization of the probe on the α sequence integrated in the mitochondrial genome (1.9- and 0.8-kb fragments on *Hae*III digests; 4.5- and 1.9-kb fragments on *Bgl*II digests) and on linearized free α-sen-DNA molecules always associated with it.

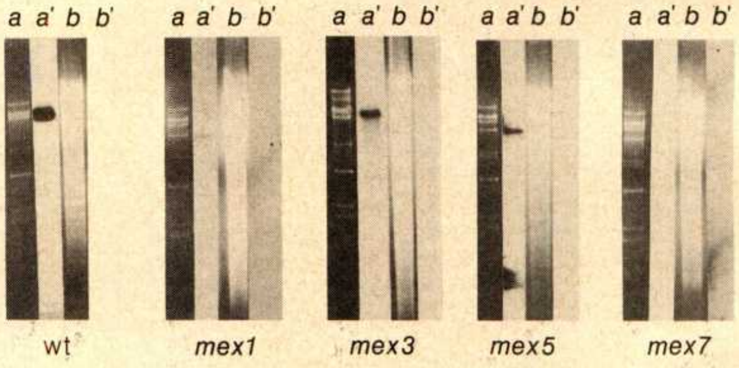


Fig. 3 Hybridization of the α sequence on the nuclear DNA of the mitochondrial long-lived mutants *mex1*, *mex3*, *mex5*, *mex7* and of *s mt+* wild-type strain (wt). The structure of the mt DNA of each mutant has already been described[14,15]: this DNA is deleted for most of the α mitochondrial intron in *mex1*, *mex5* and *mex7* mutants and it is not modified in this sequence in *mex3*. Sequencing data indicate that the *mex5* mt-DNA has retained 387 bp of the α sequence whereas only 32 bp are conserved in *mex1* and *mex7* mt-DNA. The nuclear DNA and mt-DNA of each strain were purified from total DNA extracts after DAPI–CsCl centrifugation. After digestion by *Eco*RI and gel electrophoresis, each nuclear and mt-DNA was transferred to nitrocellulose filters and hybridized with labelled α sequence (see Fig. 2 legend) as in Fig. 1. The electrophoretic patterns of the mt-DNA and of the nuclear DNA are represented in lanes *a* and *b*, respectively, and the corresponding autoradiograms in lanes *a'* and *b'*. These autoradiograms were observed after 4–5 days of exposure of the filters at −70 °C except for the *mex3* and wild-type mt-DNA (24 h at room temperature). Note that the specific activity of the probe and the hybridization conditions were such that the *Eco*RI fragment of *mex1* and *mex7* mt-DNA, which contains only 32 bp of the α sequence could be detected. In these conditions no hybridization was observed on the nuclear DNAs of these mutants.

In *Podospora*, senescence[5] is assumed to be caused by the amplification of some mitochondrial (mt) DNA sequences, called sen-DNAs[7–11]. The most frequently amplified sequence, α-sen-DNA, corresponds exactly to the first intron of the mt-gene encoding the first subunit of cytochrome oxidase[12,13]. We looked for integration of this α sequence into the nuclear genome of three senescent cultures of the wild-type strain *smt+* and in that of four long-lived mitochondrial mutants (*mex1, mex3, mex5* and *mex7*)[14,15]. Note that *mex1*, isolated in our laboratory[10], was reported by Wright and Cummings[4] to contain the α sequence integrated within its nuclear genome.

The nuclear DNA of these different strains was hybridized with the α sequence (Fig. 2) in conditions allowing us to detect a single copy of a nuclear sequence (see Fig. 1 legend). To exclude the possibility that the putative labelled bands could result from hybridization of the probe on contaminating DNAs (mt-DNA or free α-sen-DNA molecules which are found as bands between those of nuclear and mt-DNA in a buoyant density caesium chloride gradient[6,7]), hybridizations were simultaneously performed on mt-DNA and purified α-sen-DNA. Figure 2 shows the results of the hybridization of α sequence on the *Hae*III- and *Bgl*II-restricted nuclear DNA from three independent senescent cultures, A, B and C. Both enzymes have one site in the α sequence. Consequently, we expected to find two or three labelled bands if the α sequence is integrated in the nuclear genome as a single copy or tandem multicopy,

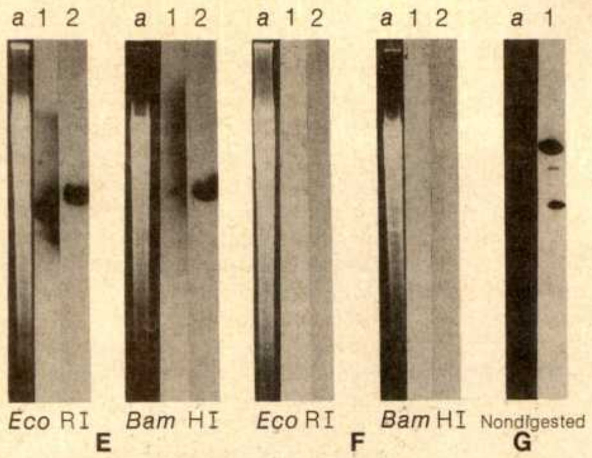


Fig. 4 Hybridization of the purified α sequence and pBR322 DNA on two nuclear DNAs obtained from two independent cultures of the *mex1* mutant. Each nuclear DNA was digested by *Eco*RI and *Bam*HI, transferred to nitrocellulose filter and hybridized with purified α sequence (see Fig. 2) and separately with pBR322 DNA (extracted according to Maniatis *et al.*[20]) used as radioactive probes. For each nuclear DNA samples (E and F), the electrophoretic patterns of the *Bam*HI and *Eco*RI digestions are represented in lanes *a*. The autoradiograms (4–5 days at −70 °C) observed after hybridization of the α sequence and pBR322 DNA are represented in lanes 1 and 2, respectively. On the nuclear DNA E, no hybridization is observed. On the nuclear DNA D, one labelled band is observed with both probes at the same position whatever the enzymes used (*Bam*HI or *Eco*RI). This suggests that these labelled bands result from hybridization of pBR322 or derivative DNA contaminating both the nuclear DNA E and the 'purified' α sequence. Hybridization between pBR322 DNA and purified α sequence, represented in G, confirmed that this purified DNA is contaminated by pBR322. Nondigested pBR322 DNA was electrophoresed, transferred to nitrocellulose and hybridized with the purified α sequence. The electrophoretic pattern of nondigested pBR322 DNA is presented in lane *a* and the corresponding autoradiogram in lane 1. The three forms of pBR322 DNA are detected by the probe.

respectively. In the latter case, one of these bands must be 2.5 kilobases (kb) long as this is the length of the α sequence.

For cultures A and B, we observed a single 2.5-kb labelled band, which we interpret as hybridization on free α-sen-DNA molecules contaminating the nuclear DNAs. For culture C, three labelled bands are observed, one at 2.5 kb, suggesting that head-to-tail copies of α are integrated, and the others, which would correspond to the sequences flanking α, migrate as the mt-fragments that hybridize with the α sequence. Our conclusion, confirmed by using other enzymes (data not shown), is that this nuclear DNA is contaminated both by mt-DNA and by free α-sen-DNA. Such contamination is difficult to avoid, even after three or four cycles of purification on CsCl gradients, because of the small differences in density of the three molecular species recovered from senescent cultures.

Figure 3 shows the results of hybridization of *Eco*RI-restricted nuclear and mt-DNAs of the mutants *mex1, mex3, mex5* and *mex7* with the α sequence which has no *Eco*RI site. Nuclear DNAs did not hybridize with the probe.

Figure 4 represents hybridization of α and pBR322 DNA probes on two different nuclear DNA extracts (E, F) of *mex1*. We found one labelled band at the same position with both probes on nuclear DNA (E) but no label on nuclear DNA (F). The nuclear DNA (E) was probably contaminated by pBR322 or derivative DNA. Contamination by plasmid DNA is fairly common; Shmookleir Reis *et al.*[16] have pointed out the danger of systematic microbial contamination of DNA samples. Indeed, the presence of plasmidic DNA, both in the DNA used as a radioactive probe (even after two cycles of electrophoretic purification[17]) and in the DNA samples tested, led to the appearance of labelled bands that could be misinterpreted in the absence of appropriate controls such as hybridization using the vector alone as probe.

Our experiments, repeated several times, failed to demonstrate any similarity between the α sequence and the nuclear DNA, from either senescent cultures or long-lived mutants. This contradicts the results of Wright and Cummings[4], who interpreted their hybridization experiments, for senescent cultures as well as for *mex1*, as positive. Two hypotheses can be proposed to explain these discrepancies: (1) the nuclear DNA from the subclones studied by Wright and Cummings contains mitochondrial sequences whereas the nuclear DNA from the subclones we have tested does not; and (2) whichever subclones studied, the nuclear DNA has no similarity with the α-mitochondrial sequence, the additional controls included in our experiments explaining their results as being caused by contamination by DNA from mitochondria or plasmids. In the absence of sequencing data demonstrating that a nuclear DNA actually flanks the α sequence, we favour the second hypothesis.

Our data show that a transposition of α sequence to the nucleus is not systematically associated with the senescence process and therefore cannot explain the arrest of growth in *Podospora*. This conclusion is also consistent with the genetic data[18] collected previously which shows that the senescent state is maternally inherited.

I thank L. Belcour and C. Vierny for help and advice, J. Beisson and L. Sperling for critical reading of the manuscript; A. Sainsard and O. Begel for help; and Y. Brygoo and R. Debuchy for the gift of the plasmid containing the *su4* nuclear gene of *Podospora*. I acknowledge the receipt of an ADRC fellowship. This work was supported by a grant from the Centre National de la Recherche Scientifique (ATP Microbiologie).

Received 3 September; accepted 28 October 1986.

1. Farelly, F. & Butow, R. A. *Nature* **301,** 296–301 (1983).
2. van den Boogaart, P., Samello, J. & Agsterribb, E. *Nature* **298,** 187–189 (1982).
3. Jacobs, H. T. *et al. J. molec. Biol.* **165,** 609–632 (1983).
4. Wright, R. M. & Cummings, D. J. *Nature* **302,** 86–88 (1983).
5. Rizet, G. *C.r. hebd. Séanc. Acad. Sci. Paris* **237,** 838–840 (1953).
6. Stahl, U., Lemke, P. A., Tudzynski, P., Kück, U. & Esser, K. *Molec. gen. Genet.* **162,** 341–343 (1978).
7. Cummings, D. J., Belcour, L. & Grandchamp, C. *Molec. gen. Genet.* **171,** 239–250 (1979).
8. Jamet-Vierny, C., Begel, O. & Belcour, L. *Cell* **21,** 189–194 (1980).
9. Cummings, D. J., Laping, J. L. & Nolan, P. E. in *The Organisation and Expression of the Mitochondrial Genome* (eds Kroon, A. M. & Saccone, C.) 97–102 (Elsevier, Amsterdam, 1980).
10. Vierny, C., Keller, A. M., Begel, O. & Belcour, L. *Nature* **297,** 157–159 (1982).
11. Koll, F., Begel, O., Keller, A. M., Vierny, C. & Belcour, L. *Curr. Genet.* **8,** 127–134 (1984).
12. Osiewacz, H. D. & Esser, K. *Curr. Genet.* **8,** 299–305 (1984).
13. Jamet-Vierny, C., Begel, O. & Belcour, L. *Eur. J. Biochem.* **143,** 389–394 (1984).
14. Belcour, L. & Vierny, C. *EMBO J.* **5,** 609–614 (1986).
15. Koll, F., Vierny, C. & Belcour, L. *Plasmid* **14,** 106–117 (1985).
16. Shmookler Reis, R. J., Lumpking Jr, C. K., McGill, J. R., Riabowol, K. T. & Goldstein, S. *Nature* **316,** 167 (1985).
17. Barbacid, M. *Science* **225,** 670 (1984).
18. Marcou, D. *Ann. sci. natn. bot.* **11,** 653–764 (1961).
19. Debuchy, R. & Brygoo, Y. *EMBO J.* **4,** 3553–3556 (1985).
20. Rigby, P. W., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113,** 237–251 (1977).
21. Maniatis, T., Fritsch, E. & Sambrook, J. *Molecule Cloning; a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
22. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1985).

Thermal erosion by komatiites at Kambalda

GROVES *et al.*[1] have added a further twist to the continuing debate on thermal erosion by komatiite magma. This concept was introduced by Huppert *et al.*[2,3] who argue that hot, low-viscosity komatiite lava flows turbulently and transmits heat efficiently to floor rocks. When the flow rate is high enough and flowage continues long enough, floor rocks melt and are assimilated by the komatiite.

Groves *et al.*[1] present geological observations supporting thermal erosion, but claim the process operates only when a sulphide layer lies at the base of a flow between komatiite lava and floor rocks. They suggest that the sulphide transports heat from the komatiite to the substrate because of its high thermal conductivity, thereby promoting melting. I believe this concept is implausible on fluid dynamic grounds, and that some of the geological observations of Groves *et al.*[1] are open to alternative interpretation.

Figure 1*a* shows the temperature distribution in turbulently-flowing komatiite lava. The temperature remains constant nearly to the lower contact with the country rock, and decreases only in a thin boundary layer. Heat from the 1600 °C komatiite is transported directly to the floor rocks, causing them to melt. The effect of interposing a layer of any other material between the komatiite and the country rock substrate (Fig. 1*b*), even a layer of highly conductive sulphide, is to decrease the rate of heat transport, and to diminish the likelihood of melting. Using parameters listed by Huppert and Sparks[3], and assuming a conductivity for molten sulphide of 10 $\text{W m}^{-1}\text{ K}^{-1}$, I calculate that a 1 m thick sulphide layer decreases heat flux by around two orders of magnitude, if heat transport is solely by convection. Since the sulphide layer is heated from above, thermal convection is unimportant; if flowing komatiite stirs the sulphide, its insulating effect is diminished, but not eliminated.

The apparent association between erosion features and sulphide[1] may result because dense, highly fluid sulphide would naturally accumulate in depressions in the base of the flow, including any formed by thermal erosion.

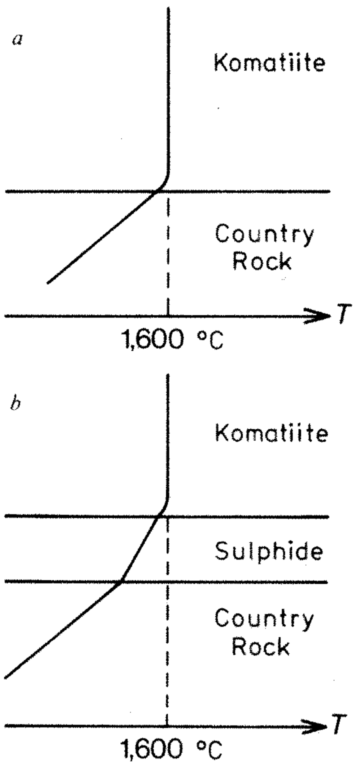


Fig. 1 *a*, Temperature (*T*) distribution in a komatiite flow and underlying rock and *b*, the insulating effect of an intervening sulphide layer.

The komatiite flows envisaged by Groves *et al.*[1] are not lethargic, merely capricious. These authors present convincing evidence that the second komatiite flow in the Kambalda sequence eroded 1 m from the top of the first komatiite flow. For komatiite to be eroded, it must be heated to 1400-1500 °C. Only then is it sufficiently molten to be removed by overriding lava. A lava flow capable of eroding komatiite would erode basalt (which melts at 1200 °C) three times faster, and sediment (which melts at 900 °C) seven times faster[2]. The lowermost komatiite flow at Kambalda, which is thicker and probably more magnesian than its successor[4], should have been capable of producing deep troughs in its basaltic and sedimentary substrate, not merely "second-order irregularities"[1].

Most of the geological features that lead Groves *et al.*[1] to reject the idea of extensive thermal erosion can be explained if we remember that volcanic eruptions are not continuous and that lava does not always follow the same course, but meanders, dams up, and spills out in different directions as the eruption rate fluctuates. Normally eruption rate climbs to a maximum, then declines. At some stage during the decline, the komatiite becomes incapable of erosion. Note also that erosion is not instantaneous but proceeds at a finite rate. Preservation of sediment beneath a komatiite flow may therefore indicate that erosion ceased because the flow rate declined, or the course of the komatiite lava changed. Erupting komatiite initially follows pre-existing valleys, and some primary topographic control is likely. The komatiite flow deepens these valleys, and flowage is concentrated in central feeder channels; a contrast in texture and composition between feeders and flanking units is to be expected. When komatiite encounters a permeable, water-saturated horizon such as flow-top or pillow breccia, hydrothermal circulation removes heat from the komatiite, and the erosion rate decreases. There is an enhanced probability of erosion ceasing when the komatiite reaches a permeable water saturated horizon, and this may appear to be caused by stratigraphic control.

Thermal erosion by komatiite flows is a new concept which may explain some of the geological and chemical features of komatiites that have puzzled us in the past. The task now is to return to mines like Kambalda, or to other areas where komatiites are well exposed to collect new petrologic and chemical data, and to use this information to test the concept. Groves *et al.*[1] have done this in their investigation of interspinifex sulphides, but other aspects of Kambalda geology seem open to alternative interpretation.

N. T. ARNDT
Max-Planck Institute für Chemie,
Postfach 3060,
D-6500 Mainz, FGR

1. Groves, D. I., Korkiakoski, E. A., McNaughton, N. J., Lesher, C. J. & Cowden, A. *Nature* **319**, 136-139 (1986).
2. Huppert, H. E., Sparks, R. S. J., Turner, J. S. & Arndt, N. T. *Nature* **309**, 19-23 (1984).
3. Huppert, H. E. & Sparks, R. S. J. *J. Petrol.* **26**, 694-725 (1986).
4. Lesher, C. M. thesis, Univ. Western Australia (1983).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):
☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify)______
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name______
Title______
Institution/Company______
Department______
Address______
City______ State______ Zip______
Phone (____) ______

UMI Article Clearinghouse

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

nature

How to give good dogs bad names

The British nuclear industry has been given a year to clean up its leaky separation plant, but it will take longer to repair the damage the leaks have done to the reputation of nuclear power.

THERE is nothing quite like the nuclear industry for fossilizing the technology of the past. A nuclear installation, once commissioned, becomes radioactive and thus inaccessible by people, with the result that radical improvements of performance are virtually impossible. This simple truth lies at the root of last week's fierce criticism, by Britain's Nuclear Installations Inspectorate, of the reprocessing plant at Windscale (now called Sellafield) in Cumbria, on the north-east seaboard of England (see p. 603). The plant was designed in a hurry nearly 40 years ago, soon after Britain decided to use plutonium for nuclear bombs. The plant reprocessing the bulk of spent nuclear fuel from British nuclear reactors was built nearly 30 years ago, and afterwards extended in bits and pieces. Throughout the life of the plant, heroic efforts have been made to keep it working safely and sucessfully, but they have not prevented a long string of embarrassing mishaps that would have seemed comic were they not also potentially tragic. Now the nuclear inspectors have rightly said that this must stop, and the Health and Safety Executive has used its legal powers to tell British Nuclear Fuels (BNFL), the publicly owned company that operates the plant, that it has a year in which to put its house in order. "High time, too", will be the general reaction.

So it is, but for a variety of reasons not all as simple as the evident and understandable exasperation of the nuclear inspectors' report suggests. It goes without saying that the mishaps of technical incompetence belie the now-forgotten boast of British nuclear engineers in the 1950s that their techniques would set standards of excellence which, when followed by the rest of British industry, would turn Britain into a temple of high technology. At worst, while accidental leaks of radioactive materials continue, fears will persist that some future mishap will expose large numbers of people to potentially damaging doses of radiation. So it is also proper to acknowledge that, so far, the Sellafield reprocessing plant has not been a significant hazard to public life and health. In part, the reputation of the plant has been sullied by its proximity to the Windscale plutonium-producing reactor, where a fire in 1957 caused the world's most serious release of artificial radioactivity until the Chernobyl accident on 26 April this year. By comparison, the effects of Sellafield's string of mishaps are trivially unimportant.

Safety

Britain's nuclear inspectors conclude that a plant whose safety should be unquestioned cannot be allowed regularly to advertise its frailty. But may not the cumulative effect of Windscale's transgressions be more serious? That is the question that environmentalists are forever asking. There are three respects in which the question is apt and deserves a patient answer. The government of the Republic of Ireland is probably right to complain that the Irish Sea is more heavily contaminated with radioactivity than any other stretch of the world's continental shelves, but BNFL is probably also right to say that the contamination of human food chains is within acceptable limits; even though the discharge is now to stop, continued monitoring will be needed for decades to come. The occurrence of clusters of childhood leukaemia in villages near the plant is another puzzle; a report by Sir Douglas Black two years ago may have correctly concluded than the clusters could be statistical fluctuations, but a diligent search for more immediate causes is plainly necessary. Finally, while Sellafield has a good record of keeping individual workers' radiation doses within statutory limits, the annual collective dose to Sellafield workers is uncomfortably large.

What else is there to complain about? The inspectors inevitably and, up to a point, fairly direct their chief criticisms at BNFL and its managers at Sellafield. But this is where the complications begin. In the old days the operators of the plant enjoyed a reputation for needless secrecy excused, but never fully justified, by the uses to which their chief product, plutonium, may be put. They were also in those days complacent, believing themselves to be fully in charge of a technically advanced plant (which was also true). But the inspectors underestimate the traumatic effects of the period, in the early 1970s, when the stuffing was knocked out of the British nuclear industry by the coincidence of several misfortunes, chiefly government disenchantment with nuclear power (born of the penury in which British governments have rubbed along ever since, but strengthened by the industry's over-optimism about its then new advanced gas-cooled reactor) and growing public hostility. The most damaging complaint in the report is that the present managers at Sellafield have taken to substituting formal approvals by advisory committees for their own judgements, thereby giving bureaucratic processes precedence over practical solutions for real problems. The old arrogance has given way to timorousness, no better suited to the operation of a potentially dangerous plant.

Future

In the present demoralized state of British public life in which nothing much is thought possible, the prospects are not bright for the year that BNFL has been given in which to put Sellafield in order. No doubt the company can fire a few managers and hire some others, but there is little hope of being able physically to rid Sellafield of its huge inventories of unwanted fission products and contaminated material until another public agency called NIREX (Nuclear Industry Radioactive Waste Executive) can negotiate a few sites at which waste materials might be safely stored. BNFL plans to start vitrifying highly active wastes roughly a year from now, but has nowhere other than Sellafield to house the products. And some of the unused facilities that clutter the overcrowded site cannot be dismantled because they are safer as intact relics than if they were broken into smaller pieces. Can NIREX, after five years of failing to find disposal sites, work magic in time to let BNFL complete its assignment?

The inspectors also complain that BNFL has been concerned with the short-term tactics of survival to the exclusion of strategic planning, which may be another consequence of Sellafield's recurring mishaps: BNFL's board members have been made into public relations men (which is not usually their talent). But much of the trouble is external to the British nuclear industry. In 1975, when BNFL had planned a new uranium oxide reprocessing plant at overcrowded Sellafield, it believed it could not follow the sensible course of putting the plant on a second site for fear of a bruising public inquiry and, possibly, refusal. Ironically, it was forced to suffer the inquiry anyway. In present circumstances, the radical strategy of abandoning the old reprocessing plant, replacing it with another, might be safer and even cheaper, but the political obstacles to such a course would be formidable if not insurmountable.

So who is really to blame? The oscillation of successive British governments' policies on nuclear power are largely responsible for the lack of timely investment at Sellafield. The present government, while mildly supportive of nuclear energy in the sense of being prepared to back plans for building new reactors has shrunk from its more onerous responsibility, that of winning by persuasion a civilized accommodation with those dissenting from its own policies, which is the hallmark of good government. But what government would take up cudgels on behalf of an organization as accident-prone as BNFL's Sellafield? In the long run, BNFL's most serious failing may be none of those listed by the inspectors last week, but that of having provided the means by which British nuclear power has been made a formal political issue, with the Labour opposition party committed to the 'phasing out' of nuclear energy. □

Slogans for recovery

The United States must improve its industrial performance in its schools and colleges.

If the Iranian arms scandal blows over soon, some attention may be given in Washington to the emergent slogan of competitiveness, the means by which the US government and its advisers are seeking to accelerate the changing pattern of trade between the United States and the rest of the world. The motives are clear and even admirable. In spite of the devaluation of the US dollar relative to the yen and the deutschemark during the past year, the US trade deficit has not shrunk but, rather, grown to a larger amount than ever, $37,700 million in the quarter ending in September, nearly 6 per cent more than in the previous three months. Although the volume of imports to the United States has indeed shrunk, higher prices (in dollars) of imported goods have kept the trade deficit growing. So it makes sense that the authorities should urge US industry to be more competitive.

Will the new bout of exhortation be more effective than the great fuss about declining US productivity three years ago? There are some obvious snags. In the fields in which US pride has recently been most humbled, motor cars and microchips, there is no reason to believe that US manufacturers invest less in capital equipment and automation than their more successful Japanese competitors. Nor is the United States short of the capacity for engineering innovation. The defects are more subtle — and thus will take longer to repair. The history of General Motors in the past few months is an illustration of how the managements of corporations can become detached from the products they make and from their ultimate customers. And while most of the new ground in microcircuit design has been broken by US companies, the constant worry about quality (or reproducibility) owes much to the shortage of skill on the production lines. Then the development of new products in typical US corporations has been bureaucratized to a degree that must delight the Japanese.

This is why the grave committees now being assembled to brood about competitiveness are certain to conclude that one ingredient of any long-term cure must be a means of persuading many more skilled people into manufacturing industry. Countries such as Britain have long sought solutions to this problem, but without success so far. But even if everybody knows what needs to be done, and has the will to do it, radical improvements of the educational system take decades, not months. Yet the only short-term solutions are crude and unpalatable — reductions of lving standards. That is what the accumulations of trade deficits will, in the end, dictate. □

Solstitial considerations

The disappointments of 1986 have been balanced by good tidings of many kinds.

Before readers of this and other periodicals lapse into the torpor of what seems increasingly to be an international holiday, with elevator muzak playing Christmas carols even in Japan, 1986 should not be mistaken for an entirely disappointing year. True, many of the disappointments have been cruel — Chernobyl, for example. AIDS came to seem more serious a threat to public health than had been thought. The summit meeting meant merely as a preparation for a summit raised peoples' hopes for arms control and then dashed them. The pace of economic growth has been disappointing, especially after the fall of the price of oil (which now seems to be on the way up again). Most countries in the West became more despondent about the difficulties of sufficiently educating the young. There was more than the usual harvest of discovered scientific fraud. Although Orlov was able to leave his Siberian labour camp, his comrade Anatoli Marchenko died in prison only a week ago. The apparatus of the US presidency has once again behaved as if above the law, and seems prepared to trade for hostages as well.

If the good news is, as always, harder to find; it is not less important on that account. People make their own tunes when whistling to keep up their spirits; for some, it is sufficient that there has been no final cataclysm, but that is too little to ask of a dead year, especially one in which there were cheerful signs of beneficial developments. Researchers, pilloried as they often are as denizens of ivory towers, may take particular pride at the speed at which some laboratory studies have been turned to practical ends. Chernobyl, for example, was heroically contained and has been understood. The fear that a large proportion of those infected by the AIDS virus may develop the disease has been partly offset by the hope that some kind of prophylaxis will emerge from the energetic attempts to develop a vaccine.

When it seems that changes of atmospheric ozone in the Antarctic spring may be sensitive indicators of anthropogenic threats to the ozone layer as a whole, the US Navy and the National Science Foundation are able to carry out speedy measurements, suggesting that growing fears of global change may at least be matched by improvements of monitoring capacity (and of understanding). Parallel computers have at last come of age (or will do so when programs have been written for them) while there is new generation of semiconductor devices, formed from GaAs by epitaxial growth, waiting in the wings.

In basic science, 1986 has been distinguished by its ingenuity, typified by the scanning electron microscope (which belongs to an earlier year, but whose inventors were awarded a Nobel prize in that now finishing). The argument (which may be wrong) that a 60-year-old set of gravity measurements contains evidence of a previously unknown force, in the same vein, only helps to show what should never have been forgotten, that the big battalions manning the accelerators do not have all the ammunition. The quasi-crystal saga has provoked many new lines of inquiry into the true structure of solids, while chemists must seriously think again about the electronic structure of molecules with conjugated bonds. The *in vitro* manipulation of plant and animal genes by people is now exquisitely delicate, but the real-life manipulation of genes by protein regulators has begun to be understood. It is no surprise that all this (and much more) should have happened in an otherwise unspectacular year; the surprise is merely that it has happened at a time when budgets have generally been cut, which may say something important about the response of researchers to adversity. The pity is that all this excitement has not been adequately communicated to those who ultimately, through taxes, pay the bills, as well as to their children. For 1986 seems also to have been the year for realizing that too few young people are being adequately trained to keep all these torches alight. □

Nuclear reprocessing

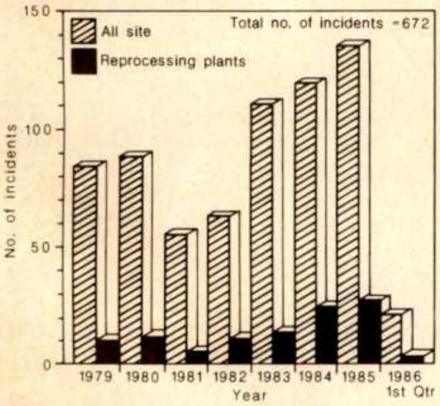
Sellafield threatened with closure within a year

THE management of Britain's only nuclear reprocessing plant, at Sellafield in Cumbria, has been given a year to improve the safety of the site or face closure.

The stern warning has been issued by the Health and Safety Executive (HSE), the licensing authority for the site, after an audit which it describes as "one of the most thoroughgoing physical examinations ever carried out in the UK at one time on any complex plant". The management of the complex, now operated by British Nuclear Fuels Limited (BNFL), is under no illusions about the seriousness of the findings, says HSE.

According to Mr John Rimington, HSE's director general, last week: "Our findings are not recommendations; they will have to be implemented. They match the comprehensive nature of the review. We intend that they should be carried out without delay".

HSE says that the BNFL management has attempted to improve the safety of the processing plant, which is on the site of Britain's oldest nuclear power station. But since 1979, there have been 672 'incidents' at Sellafield. The improvements have not been enough, concludes HSE. According to Rimington: "The main policies and priorities pursued by BNFL in improving the plant seem to us to have been the right ones. But the application of these priorities has reduced the attention paid to potentially hazardous parts of the older plant. Some of these do not yet come up to the scrupulously high standard we demand of the nuclear industry."



One of the most 'hazardous' parts of the plant worrying HSE is the central area/building which performs the initial reprocessing, B205. The executive intends to put political pressure on BNFL to conform and Sellafield's operating licence will be amended to ensure that the improvements are implemented.

HSE gave particular attention to the five incidents this year which caused disquiet and, in part, provoked the audit:

- 23/24 January — discharge of uranium into the sea;
- 5 February — leak from the central reprocessing complex B205;
- 13 February — fire on site;
- 18 February — pipe leak;
- 1 March — spillage of plutonium fuel pellets and subsequent contamination of contact surfaces.

The reprocessing complex at Sellafield is the home for more than 90 per cent of the radioactive waste produced by non-military power generation in the past 30 years. There are 16 reactors in 11 Magnox nuclear power stations in the United Kingdom. The uranium fuel rods of these reactors, enclosed in magnesium alloy containers or cans (Magnox), are stored in cooling ponds at the power stations before transportation to Sellafield; both rods and containers are subject to corrosion.

The chemical separation of the re-usable elements of uranium and plutonium leaves high-level radioactive waste that is stored on site. HSE concluded that as long as this waste remains in liquid form, it can leak and "presents a hazard to those who work there, and potentially to the public or the environment". Less active low-level waste is discharged into the sea or atmosphere or is buried at the Drigg site nearby. **Bill Johnstone**

Chernobyl

Polish fallout underestimated

POLAND's official report on the Chernobyl disaster, published in a limited edition a few weeks ago, makes no mention of the consistent under-reporting of radioactive iodine contamination during the first few weeks after the accident. Nor has the International Atomic Energy Agency (IAEA) or the World Health Organization (WHO) been notified of the error, which is therefore perpetuated by the fallout maps produced by those organizations. The story has, however, been reported by several underground bulletins, notably Wokopach, the newsletter of pro-Solidarity nuclear scientists. The initial error was a genuine one. The emergency on 29 April, was originally set up in the 1960s to deal with the aftermath of a possible nuclear war, and its outdated procedures for iodine estimation systematically lost up to 75 per cent of the iodine before a reading was obtained. Normally, only a few of the 150 monitoring points operated by the public health service are working at any one time, and as the radio-iodine level was normally close to zero, the procedural error was not observed.

Early in May, scientists carrying out independent tests of fallout noted that their readings were consistently 3–4 times higher than the official figures notified to IAEA. The only exception was the official reading from Katowice, where the public health scientists had access to a new, sophisticated estimation apparatus from the United States. At the end of May, the public health network notified the error to the special commission dealing with the effects of Chernobyl in Poland.

On learning of the error, the commission decided not to make it public. It justified this decision to its scientific consultants by pointing out that, a month after the accident, the worst danger from radioactive iodine was over (at that time, the commission did not know that emissions from Chernobyl had continued until 6 May), and that to reveal the error would spread unnecessary public alarm. The danger-level for milk had been set at 1,000 Bq, corresponding, at worst, to a true value of 4,000 Bq — and the level set for Vienna was 5,000 Bq.

Whether the commission was justified in keeping the information from the Polish public is a moot point. Withholding the information from IAEA and WHO, however, is much less excusable as both these bodies would surely have appreciated the commission's evident embarassment over the error and the fear that its revelation might cause panic. Even less understandable is that the story of the error was suppressed in the official report, a publication for official distribution only. *Nature's* attention was first drawn to the matter by a government press spokesman, Mr Jerzy Urban. **Vera Rich**

European Communities

Framework programme in limbo

DESPITE arguing until 4 a.m. on Wednesday last week, Europe's science ministers have failed to agree on a budget for the European Commission's 'framework' programme of research and development, a programme that aims to lay the basis for European cooperation in key areas of research and technology up to 1991. The ministers decided only one thing; to postpone a decision on financing until 22 December, when they will meet again in the hope of completing a compromise which began to emerge in the small hours last Wednesday.

However, the new deal, which involves a three-year rather than five-year programme, has a long way to go before it can form the basis of an agreement. The Commission had tabled a 7,735-million-ECU (£5,400-million), five-year programme, but last week the British, French and West German delegations were sticking on low figures of 3,500–4,000 million ECU, the Dutch on 6,000 million and the other eight delegations on the full amount proposed by the Commission. The gap between the extremes was thus more than a factor of two, and it was difficult to see last week how it could be bridged by the Commission's compromise proposal, of 3,700 million ECU to be committed, and 3,100 million ECU to be spent. These sums would cover three years not five.

Also against the new proposal is the European Parliament, whose energy, research and technology committee has condemned any reduction in the programme, and urged the Commission to prepare to withdraw the whole proposal. Jacques Delors, the president of the Commission, in reply to an address by the British Prime Minister, Mrs Margaret Thatcher to the parliament in Strasbourg last Tuesday, warned that he had instructed research commissioner Karl-Heinz Narjes to abandon the framework programme if ministers failed to give it a reasonable level of financing. Thatcher had just addressed the parliamentarians on the six months of British leadership of the European Communities (as president of the decision-making Council of Ministers) without making any serious mention of research. Delors gives the framework programme a high priority, and was clearly angered that Thatcher had proved an obstacle to its implementation. However, within a few hours, Narjes was proposing the new compromise. Parliamentarians have since expressed their dismay at this outcome, and are pressing for a thoroughgoing confrontation between the Commission and the Council of Ministers.

Finally, however, it may be a smaller straw that breaks the Commission's back: the matter of a 12.6-million ECU increase in the 700-million-ECU four-year budget of the Joint Research Centre's, the group of four laboratories in Italy (Ispra), West Germany (Karlsruhe), Belgium (Geel) and the Netherlands (Petten) which are wholly funded by the Commission. Last week the research ministers received a critical report on the centres prepared by an independent group of European industrialists, and although the Commission already has plans to travel down the road the industrialists recommended (greater interaction with industry, better management of funds), most ministers were sufficiently confirmed in their negative view of the centres that they refused any increase. The exception was the Italian minister, who benefits from the considerable research employment that the largest of the laboratories, Ispra, provides.

Previously fully in support of the framework programme, the Italian delegation has now linked its support to agreement on the increase for the research centres. This must make considerably less likely the unanimous agreement that is necessary for the framework programme to go forward, and it seems probable, barring a miracle, that the framework programme will limp forward, unagreed, to the next six months' Belgian presidency of the Council of Ministers. This would leave a number of programmes hanging loose, such as ESPRIT in information technology, RACE on the development of Europewide integrated broad-band communications, EURAM on advanced materials and the unstarted joint research programme on AIDS. There is around 1,050 million ECU in the European kitty next year to continue existing programmes, but this is not enough to increase the pace of cooperation intended.

Robert Walgate

Embryo research

UK legislation further delayed

STILL wriggling uncomfortably on the hook of human embryo research, the UK government last week issued a consultation paper on *Legislation on Human Infertility Services and Embryo Research* (Cm 46, HMSO, £2.90), asking for responses by the middle of 1987. The aim is then to introduce legislation, although it seems likely that a general election will intervene.

The paper is based on the Warnock report, published in 1984, since when the government has avoided three attempts by members of parliament (MPs) to prohibit almost all human embryo research, while promising its own legislation. From these attempts and other responses, the paper divides Warnock's recommendations into those where "a broad measure of agreement is likely" and those where it is not.

Chief among the contentious issues, according to the paper, is the use of human embryos for research purposes. One camp argues that research is both necessary and warranted. It is necessary to improve infertility treatments, contraceptive techniques and the detection of early abnormalities. And it is warranted for 14 days after fertilization. The other camp is opposed to any research that leads to the destruction of an embryo because this "is tantamount to murder", says the paper.

Pointing out that even the members of the Warnock committee were split on this issue — as are government ministers — and that various intermediate positions have been expressed, the paper says that the government intends to offer parliament alternative sets of draft clauses dealing with the issue, and to allow MPs to vote according to their consciences rather than along party political lines.

Other issues that are considered contentious are the possibility of egg or embryo donation and whether human sperm should ever be allowed to fertilize animal eggs, even as a test for sperm quality with destruction of the embryo after one division.

Although commercial surrogacy was made illegal in 1985, the paper invited comment on whether the law should be amended to allow the surrogate mother and the commissioning parents to be prosecuted. It also seeks views on the possibility of a non-commercial surrogacy service.

Only one reaction to the consultation paper is unanimous, namely that it is a delaying tactic. Mrs Phyllis Bowman, director of the Society for the Protection of Unborn Children, says the paper is nothing but a rehash of the Warnock report. A spokeswoman for Progress, a lobbying group in favour of human embryo research agrees but sees the delay swinging public opinion in the group's favour.

The last attempt to prohibit human embryo research was on 13 October when Mr Ken Hargreaves MP introduced legislation in a way that stood no chance of being accepted but at least forced a vote. As on past occasions there was a substantial, if reduced, majority in favour of the proposed bill. Meanwhile, such research continues under the auspices of the Voluntary Licensing Authority set up last year by the Medical Research Council and the Royal College of Obstetricians and Gynaecologists. The authority tends to follow the recommendations of the Warnock report, which are sometimes more liberal than local ethics committees through which medical research proposals have also to pass.

Peter Newmark

Medical research

Backs to the wall in Britain

The British Medical Research Council (MRC) faces a deficit of £13.2 million by the end of the decade even if it does no more than maintain the existing rate of awards and renew long-term grants at an appropriate level. But there is also a pressing need to spend more on research into AIDS (acquired immune deficiency syndrome) and to offer more competitive salaries to keep the best scientists, says Sir James Gowans, secretary of MRC.

Introducing the council's annual report in the year before he retires, Gowans described AIDS as a "national emergency" that deserves special earmarked research funds. Discussions with government departments seem likely to produce the £0.5 million needed immediately for one major and two minor AIDS projects, but several million pounds will be needed over the next few years. (A bare £2 million was spent on AIDS research by MRC and the Department of Health and Social Services in the past year compared with the $40 million predicted for 1987 in the United States.) Priority must be given to monitoring the spread of the virus that causes AIDS and the effect of the campaign to reduce the spread, said Dr D. A. J. Tyrrell, chairman of the MRC working party on AIDS. But neither he nor Gowans would be drawn on how to overcome the problems that are holding up monitoring (see *Nature* **324,** 396; 1986).

On salaries, Gowans expressed frustration that MRC's grant-in-aid, worth £129 million last year, assumes an increase of only 3 per cent this year and 2.5 per cent next year when pay awards are running at 5–6 per cent. The level of awards for its staff is not controlled by MRC as it is tied to increases for their university, clinical and civil service equivalents. That, and the decreasing budget, also constrains the ability of MRC to offer salaries that compete with offers from industry and overseas. At the top level, however, the council is free to offer more competitive salaries as long as the mean is maintained at the average for professors in universities, which have the same flexibility.

Unless the MRC budget is increased, said Gowans, it will be extremely difficult for the council to grasp new opportunities in research. The problem is already very serious; 90 project grants were approved but not funded in the past financial year. Royalty income from discoveries made by MRC scientists is expected to increase in coming years, particularly through the activities of the new Centre for Collaborative Research. But last year it was only £0.1 million and a limit of £2.5 million has been imposed. **Peter Newmark**

Japanese volcanoes

Forecasters wise after the event

Tokyo

Japan's Volcanic Prediction Liaison Council seems to have been rewarded for its failure to forecast the behaviour of Mount Mihara on Oshima, the volcanic island south of Tokyo evacuated during a major eruption last month. The council has won substantial new facilities on the island that will make Mihara one of the best-monitored volcanoes in the world.

The council, advisor to the Meteorological Agency, has come under fire for its pronouncements about Mount Mihara. At the end of October, following the outbreak of volcanic tremors, the council declared that any eruption would be "several months or years" away. Even when the lava started spewing out of the volcano two weeks later, it was confidently predicted that the flows would be confined to the caldera. But when a new line of vents broke through the caldera rim, sending lava surging towards the main town Motomachi, and the island had to be evacuated, the council turned about and started issuing warnings that mixing of sea water and magma could lead to a phreatomagmatic eruption, as at Krakatoa.

The Mount Mihara eruption at its peak on the night of 21 November.

Now, almost a month later, with the volcano silent and the government running up huge bills to support the evacuees, the council has been under pressure to issue an all-clear. Finally, last Friday, its members adopted a "unified view" that Mount Mihara is now calm "in the short term"; the Tokyo Metropolitan Government immediately announced that the islanders would be allowed to return to Oshima between 18 and 22 December. But the council's decision was given on the precondition that the island's volcano-monitoring network be substantially upgraded before the islanders return.

The number of seismometers on the island will be almost doubled, to nineteen. Seventeen new tiltmeters will be put in place along with various other pieces of equipment costing a total of Y1,100 million (£5 million). The seismometers will be linked to telephone lines and 40 engineers of Nippon Telephone and Telegraph are racing to link the seismometers to a telemetric network by 19 December.

Preliminary analysis of the November data suggests that prediction will still be unreliable, and that details will often be discovered only after the event. Seismograph data did, however, play an important role in the evacuation, according to Professor Aramaki of Tokyo University Earthquake Research Institute, who spent the night of 21 November at a police station in Motomachi. A shift in the earthquake swarms to a position off the south-east coast of Oshima, reported by telephone to Aramaki by researchers at Tokyo University's observatory on the island, coincided with police reports of new cracks in a road near Habu, the south-east port of Oshima; acting under Aramaki's advice, police removed evacuees from Habu to the northern ports of the island.

The new equipment will be operated by Tokyo University and the Meteorological Agency as well as various other government agencies. But it remains to be seen whether they can work together. While Tokyo University researchers stuck to their posts on 21 November, as a river of lava bore down on their observatory, officials of the nearby Meteorological Agency Observatory were among the first to flee the island. Feelings between the two groups are strained. **David Swinbanks**

Pesticides

FDA told to try harder

Washington

THE US authorities are not doing enough to catch and prosecute unscrupulous food growers who may be steadily poisoning the American public through illicit use of pesticides. A study* by the General Accounting Office (GAO) concludes that the US Food and Drug Administration (FDA) is the agency most at fault.

The study shows that the proportion of tested domestically produced food is small (probably less than one per cent), and, more significantly, that over half of pesticides that may be present are not detected by any of FDA's five preferred 'multi-residue' tests. Eight out of the nine most dangerous pesticides used are not detected by these standard tests. The agency uses the more time-consuming and expensive single-residue tests only when it has reason to suspect chicanery.

The proportion of tested samples has fluctuated since 1979 between 1.8 and 4.5 per cent, but more than 5 per cent of fish samples tested over the same period contained illegal residues. FDA says these proportions are not representative of the food supply as a whole, as trouble spots are sampled more often.

But the most disturbing discovery of the report is FDA's record at preventing the sale of adulterated food and prosecuting wrongdoers. In 107 out of 179 cases in which the agency found illegal residues (over a period of more than two years), the food had already been sold by the time the result came through, and in none of the 179 cases did FDA seek criminal penalties. Because the agency lacks detention authority over domestic food, it often cannot prevent such food being exported —but it can seize adulterated imports.

FDA also lacks legal authority to impose civil penalties, a shortcoming GAO would like remedied to provide an effective deterrent. The alternative, costly criminal procedures, is rarely used. Congress may hold hearings on the agency's effectiveness at pesticide control next year, and should be interested in the idea.

Does all this mean that Americans are being poisoned? Not quite. Illegal residues may simply be pesticides applied to the wrong crop, or may exceed approved 'tolerances' by a small amount. The tolerances are supposed to have large built-in safety margins. Environmental groups say, however, that the factor of 10 sometimes used in extrapolating safe dosages from animal studies is inadequate and that tolerances are set before all the relevant data are collated. **Tim Beardsley**

Pesticides: Need to enhance FDA's ability to protect the public from illegal residues. (GAO, 1986).

UK industrial research

More stick than carrot on offer

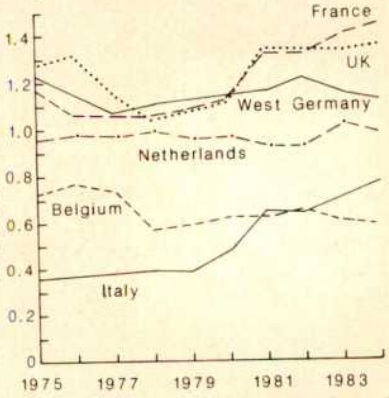
BRITISH manufacturing industry, part of which enjoys lucrative defence contracts, must itself spend more on research and development, says Mr John Fairclough, the government's chief scientific adviser.

Annual government investment in research and development is £4,300 million, of which 51 per cent is channelled into the Ministry of Defence (MOD). In sharp contrast, the Department of Trade and Industry (DTI), one of the principal sources of industrial research funds, has seen its share fall from 7.9 per cent of the pool in 1983–84 to an estimated 5.2 per cent in 1988–89.

The call for more investment came on the publication of the government's review of research and development expenditure. Although Britain invests more in publicly funded research and development than most of its European neighbours, but as a proportion of Gross Domestic Product, the total devoted to civil projects lags behind West Germany, France, Japan and the United States.

Fairclough, who began a two-year secondment from IBM (UK) to the Cabinet Office in May, is attempting to change that ratio. One solution is for companies publicly to disclose their research and development expenditure. Such disclosure, Fairclough believes, would make financiers aware of the long-term strategies being adopted by companies.

The Fairclough solution was not the only one to emerge from Whitehall last week as the government attempted to forge better links between its departments, the universities and industry.



Another, credited to the Prime Minister, Mrs Margaret Thatcher, is a programme for re-routing £210 million of the budget for scientific research and development to "speed up the development of products and services from scientific and technological research carried out by universities, government departments and industry".

The programme, called LINK, is to encourage research and development partnerships between the scientific community and industry. Government funding is spread over five years and is expected to attract the same from industry.

Mrs Thatcher, in announcing the programme, said that "the excellence of British science has not been matched by a rapid pace of application by British industry of the newly emerging technologies. Nor is the private sector investing as much of its own resources in research and development in this country as are our more successful competitors."

Fairclough — more for research.

The success of the LINK programme will require cooperation between government departments which in the past have been noted for their rivalry. They are the departments that already have research budgets — the Departments of Trade and Industry, Education and Science, Energy, Environment, Transport, Health and Social Security, Agriculture Fisheries and Food and the Ministry of Defence. The Scottish, Welsh and Northern Ireland offices will also be involved.

There will be a steering committee composed of senior industrialists, academics and officials from the relevant departments and research councils with a brief to foster strategic areas of scientific research directed towards "the development of innovative products, processes and services by industry"; to stimulate industry to increase its investment in research and development; to encourage industry to exploit developments in science; to make scientists more aware of the needs of industry; and to develop technologies that cross the normal boundaries of scientific disciplines.

The Advisory Council for Applied Research and Development, the Advisory Board for the Research Councils, the Assessment Office on Science and Technology in the Cabinet Office are collaborating to find other ways of marrying research to industry. **Bill Johnstone**

Nuclear safety

United States opts for caution

Washington

The principal US military plutonium production reactor, a graphite-moderated plant similar to that at Chernobyl, is to be shut down for six months by the Department of Energy (DoE) to carry out urgent safety modifications which could cost about $50 million.

N reactor at Hanford site.

The 23-year-old 'N' reactor on the Hanford military nuclear reservation in Washington state has been a focus of concern since the Chernobyl accident. Besides being graphite-moderated, like Chernobyl-4 it lacks a containment structure capable of withstanding steam and hydrogen explosions that might occur in an accident. Two out of six independent consultants, asked by DoE to assess N-reactor safety, recommended closing down the reactor immediately and permanently, unless to do so would threaten national security.

Others said the poor containment meant the consequences of an accident would be worse than at a commercial power station, and that they could support continuing operations only if further safety checks were made. DoE declined the invitation to shut down the reactor permanently, but has accepted the need to spend immediately $50 million on improvements.

The consultants, under Louis H. Roddis, submitted independent reports, agreed that there was virtually no possibility of a Chernobyl-like accident, but found that radioactivity might be released following a minor accident because of doubts about the N reactor's 'confinement' system, which is designed to lower pressure in the reactor building after an accident by spraying cold water. The system has never been tested under realistic conditions. One consultant commented that the list of required improvements is so extensive that a commercial reactor would be required to shut down to carry them out. Management improvements were also urged.

The consultants urgently recommended installing a hydrogen mitigation system, and a 'basin' for collecting radioactive water that would be released in an accident. Hydrogen explosions figured in both the Chernobyl and Three Mile Island accidents, but the N reactor confinement system would not withstand an explosion of the size thought possible, according to the consultants. And under present plans, contaminated water would be temporarily stored in a desert pit, a plan DoE acknowledges to be inadequate. The department says it will start to install a hydrogen mitigation system immediately and also accelerate a comprehensive probablistic risk assessment now in progress.

While announcing the temporary shutdown, DoE also said it was awarding a 5-year, $4,000-million contract to manage the entire Hanford operation to Westinghouse. The N reactor was previously managed for DoE by the General Electric Company, and some of DoE consultants noted "complacency" over safety. Another contractor at Hanford, Rockwell International, has been the focus of allegations of poor management at two plutonium production plants recently shut down temporarily for safety violations.

Whatever safety improvements are made, the N reactor will not be in business for much longer. Commissioned in 1963 with a design lifetime of 20 years, its pile is now expanding at the rate of 0.8 inches per year, putting stress on coolant pipes and threatening to hit the thermal shield in as little as 3 years. That will be the end for the reactor as replacement of the whole pipe is not feasible. DoE is now trying to decide whether it needs another plutonium production reactor (it has others at the Savannah River complex in South Carolina) for next century's nuclear weapons.

Tim Beardsley

Albanian outlook

Changing tack on exports

Ramiz Alia, who became leader of the Albania Workers' Party (AWP) on the death last year of Enver Hoxha, has an unusual approach to his country's nonrenewable natural resources. "The interests of future generations", he told the AWP congress last month, demand that such assets as copper and chromite ores be "put into economic circulation as soon as possible", by extracting and exporting them. "It is better to leave these natural assets to future generations in the form of factories, plants and combines, as means of production, as functioning assets, rather than in the form of potential resources hidden underground", he said.

Although he was referring chiefly to mineral ores, Alia's speech made it clear that Albania's new five-year plan will attempt also to increase its capacity for exporting oil and coal. Albania, the least developed country in Europe, thus differs markedly from most developing countries which see the massive export of raw materials as a vestige of colonialism, and aim instead at the export of finished or semi-finished products with a larger added value. But the policy of Albanian self-sufficiency established by Hoxha after his breaks with the Soviet Union and then with China means that the country must implement its plans for modernization without foreign loans or grants-in-aid.

In these circumstances, Alia's view that future generations can best be served by selling off the country's irreplaceable assets does make some kind of short-term sense. Albania's economy is still based largely on human muscle-power. The 'prestige' civil engineering projects of recent years, such as hydroelectric dams and railways, have largely been built by student volunteers, equipped with little more than pick and shovel. Agriculture is labour-intensive, and Alia's promise to "increase the role of science in its service" will entail not a decrease in the numbers employed but a "considerable increase".

Science in the service of production is one of the main targets of Alia's programme. This policy demands trained personnel, in a country where at present only 57 per cent of young people receive more than primary education. (By 1990, according to the plan, this figure should rise to 73 per cent.) Alia's speech stressed not only the need for more trained scientists and technicians, but also for better management of science. The Albanian Academy of Sciences, he said, must give "more initiative and thought" to the directions in which the sciences should develop, while the Committee for Science and Technology must provide "better planning, a clearer definition of priorities and more effective organization and coordination of manpower and resources". The University of Tirana and other institutions of higher education must become "bodies of consultation for the broad masses of specialists in production", while the ministries concerned must refrain from turning educational and research centres into mere "administrative apparatuses".

In spite of Alia's boast that Albania can now train all the specialists it needs, and in spite of "self-sufficiency", the country is still, of course, largely dependent on foreign scientific achievements. And to enable Albanian scientists and technicians to make use of foreign results, the Ministry of Education has been instructed to remedy the "unsatisfactory situation" in the teaching of foreign languages. **Vera Rich**

Australia

Lay head for CSIRO board

Sydney

THE initial reaction of the Australian scientific community to the appointment of former New South Wales Premier and federal president of the Australian Labour Party (ALP), Mr Neville Wran, as head of Australia's largest research body, the Commonwealth Scientific and Industrial Research Organization (CSIRO), lay somewhere between surprise and shock. The parliamentary opposition described the appointment as "appalling political cronyism". Many CSIRO scientists were dismayed by Wran's appointment, which marks the end of a 60-year tradition of having a scientist at the head of CSIRO, but this feeling seems to have given way to guarded optimism that Wran's undoubted political acumen and extensive connections with industry will be great assets.

Wran was premier of New South Wales for 10 years and federal president of ALP for 6 until his surprise resignation from both positions last July. His government would not have had to face another election until early 1988. The stated reason for his resignation was that he was tired of politics and that stepping out mid-term would give his successor time to get used to the reins of government before having to face an election.

Wran's appointment coincides with the disclosure of legislation designed to give CSIRO a new board and structure. The changes are the result of recommendations by an Australian Science and Technology Council review of the organization's operation. In the new board, policy-making and management functions have been separated to make it more like the board of a normal corporate body. The old board consisted of three full-time members, all distinguished scientists, and five part-time outside members. One of the full-time members was chairman and chief executive, a position held in the last board by Dr Keith Boardman. In the new model these posts have been split, with the chief executive becoming the only full-time member of a 10-man board. The chief executive is thus the most senior CSIRO scientist. Boardman has been appointed permanently as chief executive subject to the approval of the board which is due to meet for the first time this month.

It will be the task of the new board to align CSIRO's research ambitions with national priorities, which means increasing emphasis on industrially applicable research, particularly on that applicable to manufacturing, although Boardman is careful to point out that this does not mean a diminution in either the quantity or quality of fundamental and rural research carried out by CSIRO. In fact, the number of scientists working in rural research has increased because of more funding from outside sources despite the decreases in government money.

In addition to Wran, the part-time board members include two scientists, a farmer, a representative from the union movement and four representatives from industry, so it should be well suited to its task of linking CSIRO to the strategic needs of industry.

The Minister for Science, Mr Barry Jones, has long bemoaned the poor record of scientists for standing up in public and speaking out for their own cause, and has even gone so far as to call them "wimps". Perhaps the appointment of a politician to head CSIRO means that Jones has concluded that the best way to get the scientists' case heard is to give them a voice to speak for them.

Charles Morgan

Britain gets Cray supercomputer at last...

THE UK Science and Engineering Research Council (SERC) has at last taken delivery of its Cray X-MP/48 supercomputer, after much discussion between the British Department of Trade and Industry and the US Department of Commerce.

For months, the latter has been concerned that 'foreign aliens' might gain easy access to the supercomputer which is to be housed at the Rutherford Appleton Laboratory. The Americans wished to impose rigid restrictions on SERC, a move vigorously resisted by the British authorities.

For logistic reasons, trade officials from both sides of the Atlantic had been unable to meet to discuss the subject, so the Americans have approved the purchase to prevent any further delay but without receiving the assurances they originally sought. The trade authorities will meet soon to discuss the 'principles' of the case. The machine will be operational by the end of January and will run by SERC on behalf of the research councils and the universities. SERC says the Cray "will be a considerable boost for UK scientists and engineers engaged in a wide range of computer simulation problems", including the physics of liquids and solids, large-scale-integration circuit design and aircraft design.

The supercomputer has been acquired partly as a result of a report prepared for the Advisory Board for Research Councils, the Computer Board and the University Grants Committee, entitled *Future Facilities for Advanced Research Computing*, by a working party under the chairmanship of Professor John Forty of the University of Stirling, who has agreed to be chairman of the management committee for the supercomputer service.

Bill Johnstone

...and Indian supercomputer is on the way

New Delhi

AFTER protracted negotiations, India and the United States have signed a document paving the way for the sale of a Cray-XMP supercomputer to India. The deal was approved in principle by the US administration eight months ago but became bogged down because of disagreements over the US terms and conditions for its sale. Indian demands that its sovereign rights should not be violated in any arrangement and US fears about possible misuse by India and potential leakage of technology to the Soviet bloc were the bone of contention. The problems were resolved and mutually acceptable conditions of safeguards were finalized last week after three days of intense negotiations between the US delegation led by Assistant Secretary of State Robert Dean and officials of India's External Affairs Ministry. The signing of the safeguards document has cleared the way for India to become the first non-NATO country to receive the computer.

According to a US spokesman, the safeguards are similar to those applicable to US computers sold in Britain and West Germany. The United States, which initially insisted on unrestricted right of inspection to check diversion of technology to third countries, finally agreed to exercise the right only in the event of suspected diversion. It also wanted to station US personnel permenantly at the computer facility, but agreed to the Indian request that US participation should be a temporary arrangement to be reviewed after 2–3 years when Indians would be ready to man the facility on their own.

The supercomputer, which is expected to cost about $20 million, will be installed in New Delhi under the operational control of the Department of Science and Technology. It will be used for meteorology but the United States has agreed to let India use any spare capacity for certain stipulated other areas of research excluding nuclear or defence-related spheres.

The Indo–US deal on the supercomputer is being keenly watched by Japan. As negotiations with the United States were proving difficult, India had six months ago placed orders for three supercomputers from Japan, but the export was stopped under pressure from the United States. Now that the safeguards have been worked out, Japan will have no difficulty in exporting similar computers to India, which hopes to install four of them by next year.

K. S. Jayaraman

Rabies vaccine

Inquiry into Argentine trials

Washington

OUTRAGE and calls for investigations have followed in the wake of revelations about field trials of a recombinant rabies vaccine conducted by the Pan American Health Organization (PAHO) (see *Nature* **324**, 202, 1986). Both the National Institutes of Health (NIH) and a congressional committee have begun inquiries into the affair, in particular into the role of the Wistar Institute of Philadelphia where the vaccine was made. Although the government of Argentina has made no formal comment on the tests, members of the Argentine scientific community feel the trial was conducted without safety precautions that would have been required had the trial been carried out in the United States (see p. 610). But Hilary Koprowski, director of the Wistar Institute, says people should be rejoicing that the new vaccine seems to be a success and not complaining about an experiment now over.

The new vaccine uses a vaccinia virus genetically engineered to contain an immunogenic gene from a rabies virus. Tests on laboratory and captive animals have shown the vaccine to be successful in inducing antibodies against the rabies virus. A total of 20 cows were vaccinated at the PAHO field station in Azul, 180 miles south of Buenos Aires. Ten cows were vaccinated intramuscularly and ten subcutaneously. Each group was kept in an open paddock with 10 unvaccinated cows.

But the field test, begun in July, was terminated in early fall after a commission of Argentine scientists visited the Azul station and raised questions about the way the experiment was being conducted. The commission reported that there were no signs announcing the nature of the experiment, both wild and domestic animals could come in contact with the vaccinated cows, excreta from the vaccinated animals was not being treated and milk from the vaccinated cows was being consumed by humans and animals at the Azul station. According to Mercedes Weissenbacher, one of the commission members, some of the milk was sent to a local dairy where it was used for making powdered milk and confectionery. Koprowski says the milk should have been frozen and sent to Wistar's French collaborator Rhone-Merieux to be tested for the presence of virus.

PAHO agreed to terminate the experiment after the commission's visit to Azul. Although Koprowski complains that PAHO was prevented from obtaining further blood or tissue samples from the vaccinated animals to see if the vaccine was still effective, Weissenbacher says blood and milk samples were collected before the animals were killed last month and are being stored for analysis.

Koprowski is upset that the animals were killed before they could be challenged with virus. He maintains that initial results were successful, the virus was not shed, neither humans nor animals became sick and the vaccinated cows developed antibodies to rabies virus.

PAHO has declined to answer questions on the trial. Carlyle Guerra de Macedo, director of PAHO, admits "mistakes were made in the way the matter was handled", but says a full disclosure will have to wait until discussions with the Argentine government are concluded.

Jose La Torre, director of the Centro de Virologia Animal and a member of the Commission sent to investigate conditions at Azul, says he agrees that the vaccine trial was an important one, but he objects to the way it was conducted. Weissenbacher concurs, adding animal care workers should not have been involved in the vaccine trial without obtaining permission from Argentine health officials. In justifying performing the experiment in Argentina, a PAHO spokesman said last month that cattle losses from rabies in Argentina amounted to millions of dollars a year. But a spokeswoman for the Argentine embassy in Washington said such losses were closer to $60,000 a year.

La Torre visited Wistar in 1984 along with representatives of the Argentine pharmaceutical company Paul which was interested in trying the new rabies vaccine. Koprowski says Wistar told the Paul representative such a commercial arrangement would have to be made through Transgene, Wistar's French collaborators. Paul never contacted Transgene, but La Torre says that following his visit he wrote to officials at Wistar, informing them that they would need permission from the Ministries of Agriculture and Public Health to conduct a field trial. An Argentine law, adopted in 1903, prohibits the import of exotic viruses into Argentina. Wistar associate director Warren Cheston says a preliminary document from Wistar to Paul concerning a licence for the vaccine charged Paul with responsibility for notifying Argentine officials.

News of the Azul trial reached the Argentine scientific community when Mauricio Seigelchifer, an Argentine scientist visiting Wistar, wrote to colleagues about the experiment. Although Seigelchifer's contract with Wistar ran until the end of the year he returned to Argentina last week.

Joseph Palca

Israel

University shake-up on cards

Rehovot

A MAJOR reconstruction of the higher education system in Israel that would create a two-tier structure, halt the growth in the number of students at existing institutions and raise tuition fees has been proposed by Professor Jacob Ziv, head of the Grants and Allocation Committee of the Council for Higher Education. Ziv, who oversees the allocation of US$246 million a year to the universities, believes there is no alternative if the growing demand for higher education is to be eased.

The universities' share of the national cake has substantially declined in the past decade. In the mid-1970s, they received 7.7 per cent of total government budget allocations (exclusive of defence expenditure and debt repayment), this year it was 4.4 per cent. This reflects a change in priorities, as during the same period government support of education (below university level) and of health services went up by nearly 40 per cent.

There has been a change for the better since the present government took office in 1984, but funds remain short and there is no solution in sight for the debts of $200 million of Israeli institutions of higher education, about half of which are borne by the Hebrew University. They have become so onerous for the university that it delayed opening its school year and its president and director general have resigned.

Although Ziv, formerly dean of electrical engineering and vice president of academic affairs at the Haifa Technion, wants more government support, he also favours restructuring higher education.

His most far-reaching proposal is for a two-tier system to halt the growth of the student population at present undergoing higher education (now increasing by about 3 per cent a year) and perhaps to reduce the numbers. Young people unable to qualify for admission under the new rigid standards would go to two-year study centres, if they did well, they could eventually transfer to a university. These centres would be inexpensive "teaching factories" staffed mainly by part-time lecturers who would not enjoy the conditions or research facilities of university personnel.

A ministerial committee is now studying Ziv's proposal to raise university fees (now about $1,200 a year). But as he sees, it, less affluent students should be allowed to pay their fees after graduation.

University libraries are less well stocked than they used to be, modern equipment is often lacking and salaries are only a quarter of those paid in the United States. The result has been a brain drain and even scientists who retain their posts in Israel often do much of their research overseas.

Nechemia Meyers

Wistar's export to Argentina

SIR—Argentinian scientists have recently been disturbed by an experiment involving both humans and domestic animals infected with a live recombinant vaccine [An account of the experiment appeared in *Nature* (**324**, 202, 1986) after this letter was written]

The experiment was carried out in an experimental farm of the Panamerican Health Organization of Azul, province of Buenos Aires, following a protocol designed by the Wistar Institute of Philadelphia, it was supervised by Dr J Held (director of Centro Panamerico de Zoonosis, CEPANZO, PAHO), Ana Maria Diaz (head of the rabies unit), Elmo de la Vega (chief of pathology) and Luis Lazaro (chief of the Azul Field Station)

The experimental protocol was presented by Dr Roberto Weimann from the Wistar Institute, who flew to Buenos Aires immediately after the nature of the experiment became public, thanks to the action of a group of scientist from AADICYT, the Argentinian Association of Research Scientists Soon after, the Argentinian Ministry of Health appointed a special committee of inquiry After interviewing Weimann and visiting the premises of CEPANZO at Azul, the committee recommended the immediate interruption of the experiment, based on the following evidence

(1) The experiment was started on 1 June 1986, without the permission, consent or notification of any Argentinian human or animal health authority

(2) The effects and eventual risks of the inoculation of a rabies-vaccine recombinant virus in open-field cattle and its possible transmission to humans are completely unknown The only reports in the current scientific literature deal with laboratory animals (mice and rabbits, *Proc natn Acad Sci U S A* **81**, 7194, 1984) and foxes kept in captivity (*Nature* **322**, 373, 1986) There is particular concern about a possible alteration in the tissue localization of the recombinant virus as in the case of the Moloney murine leukaemia virus-rabies recombinant (M P Kriegler and C Perez, Scientific Report 84/85, Institute for Cancer Research, Fox Chase Cancer Center) that might produce long-term consequences difficult at present to assess

(3) There were serious violations of ethical, ecological and safety rules

- The inoculated animals were not kept in isolation, permitting the free access of other wild and domestic animals to the area, with the danger of spreading the new virus in an open ecosystem
- No warning signs were placed at the experiment area
- There were precise instructions that the same two animal handlers should milk each group of cows throughout the experiment
- The four animal handlers were not vaccinated against smallpox before starting the experiment The only sign of previous immunization considered by those responsible for the experiment was the presence of a scar compatible with vaccination
- The four animal handlers were not under medical surveillance
- Some of the milk from the inoculated cows was consumed by the animal handlers and their families without being pasteurized The bulk of the milk was sent to the local market, and distributed to the community after pasteurization
- The way in which this experiment was performed clearly infringes the World Health Organization's recommendations about experience involving recombiant DNA technology (Bethesda, 1984, Geneva, 1985)

We are not against the development of recombinant vaccine technology in view of its potential value for effective disease control But we think that recombinant vaccines, like any other recombinant DNA project involving humans, must be developed in the light of international laws and recommendations, caring ethical and ecological consequences

Moreover, we feel that our country has been illegally used as a test field for a kind of experiment that is not yet accepted in the countries where basic research on this vaccine had originated

PABLO R GRIGERA

Centro de Virologia Animal (CEVAN),
Serrano 665-1414 Capital Federal,
Argentina

- This letter is signed by 134 Argentine scientists

Defensive deterrent

SIR—At the time I was approached, I objected to the Cornell survey of National Academy members (*Nature* **324**, 4, 1986) as to the feasibility and desirability of the Strategic Defense Initiative (SDI) even though I agree that no defensive system would be able, within the next 25 years, to provide effective population defence, even if the Soviets froze their arsenals at current levels But as Reykjavik (which took place after the survey was completed) demonstrated, this is not the right question Rather, we should be asking whether, if arms control leads to a mutually agreed reduction in the offensive threat to, say, a few hundred offensive missiles, can a defensive system be effective against this much lower threat? The answer to this question is probably yes, I would guess that many of my fellow academicians who denied the feasibility of a defensive system against the current threat would concede that defence is feasible against a lower offensive threat That those responsible for the Cornell survey did not give the respondents this choice I consider to be an improperly restrictive view of the matter

Indeed, the entire episode reflects a curious bias in the academic world against the so-called defensive-dominated world A defence-dominated world in which peace is maintained by denial of aggressive intent is surely morally superior to the present offence-dominated world in which peace is maintained by threat of mutual genocide — provided the defence-dominated world and the transition thereto can be adequately stable

Rather than sanctifying the current offensive world as being the only one in which peace is maintained, academic peace studies groups ought to be examining in full seriousness the defensive world Reykjavik gives us a glimmer of hope that the first step to a defensive transition — deep reductions in offence — is not a will-o'-the-wisp How to stabilize such reduction by appropriate defensive deployment and how to redirect SDI towards this goal ought to be central elements in academics' peace agenda, much more than ill-considered, and I fear, politically-inspired, attempts to discredit any movement toward a defence-dominated world

ALVIN M WEINBERG

Institute for Energy Analysis,
Oak Ridge Associated Universities,
PO Box 117, Oak Ridge,
Tennessee 37831-0117, USA

South Africa

SIR—Britain took 96 years to get full suffrage, in four stages, the last only 58 years ago During that time, poverty (far greater than that in South Africa today) virtually disappeared, owing to the gradual increase in wealth brought about by entrepreneurs South Africa needs *more* trade and contacts with other countries so that it can continue to raise the average standard of living

The advanced countries demand that South Africa make really major social changes much more quickly than they did, at the same time they apply sanctions that cripple its power to do so Such behaviour may be explained, but not excused, by the campaign against South Africa in the media, especially television, during the past year or so The letter of J A Barnes (*Nature* **324**, 104, 1986) rather dispels the hope that an academic would show more understanding

J A D EWART

Delgany,
Solesbridge Lane,
Chorleywood,
Rickmansworth,
Hertfordshire WD3 5SW, UK

NEW PRODUCTS FOR

PLANT CELL WALL RESEARCH

- Aniline blue fluorochrome for callose
- Yariv reagents for arabinogalactan-proteins
- (1→3) (1→4)-β-glucan hydrolase: for cereal glucans

Full details supplied with each product

Further information and orders from.

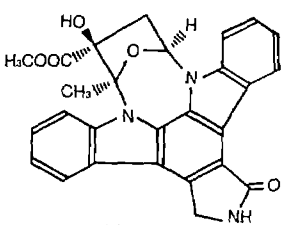
BIOSUPPLIES AUSTRALIA PTY LTD
PO Box 835
Parkville
Victoria 3052
AUSTRALIA

Telephone:
(International Access Code)-613-348 1825
Facsimile:
(International Access Code)-613-347 1071

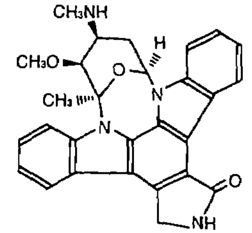
Reader Service No 9

NEW REAGENTS FOR PROTEIN KINASE RESEARCH FROM KYOWA

K-252a and Staurosporine, microbial alkaloides are potent inhibition of protein kinases with Ki of 0 7 ~ 25 nM



K-252a



Staurosporine

* Purity, >99%
* Solubility, Readily soluble in DMSO and DMF
* Stability, Stable at room temperature, store protected from light

KYOWA

To order or to request information
KYOWA HAKKO GROUP
KYOWA MEDEX CO, LTD
Ohtemachi Bldg 1-6-1 Ohtemachi Chiyoda-ku
Tokyo Japan
Tel 03 282 0095 Tlx J24543HK KYOWA
KYOWA HAKKO USA INC
New York, Tel 212 391 2604 Tlx 257105812538KHKNYK
Costa Mesa, Tel 714 556 2604 Tlx 257105812538KHKNYK
KYOWA HAKKO EUROPE GMBH
Dusseldorf Tel 0211 32 09 41 Tlx 418588026KHK D

Reader Service No 10

NEW

THE SOFTWARE DIRECTORY FOR MOLECULAR BIOLOGISTS

Christopher Rawlings

This single volume brings together the scattered information on software available to molecular biologists.

The Software Directory compares and critically appraises specialized software for molecular biology, and also provides you with concise descriptions of commercial and business software applicable to laboratory work

Each entry covers functional features, machine type, computer manufacturer, operating system, and author or company The book also contains a thorough index of all the software listed, under seven different headings (the five above, plus package name and keyword)—so that you can always find the software that is right for your work, quickly and easily

The all-new *Software Directory for Molecular Biologists* gives you accurate information on all known software in the field—commercial or non-commercial You will save time and money with this invaluable guide

300pp
$80 00
0-943818-37-0

April 1986
£40 00
0-333-39821-1

CONTENTS
Part 1—Descriptions of commercial and business software and their use in the laboratory Spread sheets, text processing, database management systems, terminal emulators, graphics, financial planning, file transfer, integrated software systems
Part 2—Directory of specialist software for molecular biologists
Part 3—Indexes and cross references of software listed
Glossary of terms

Order on FREE APPROVAL by mail or phone We accept all major credit cards

US/Canada Mail to
Stockton Press
15 East 26th St
New York, N Y 10010
Call toll-free 800-221-2123
In N Y, 212-481-1334

UK/Rest of World Mail to
Elaine Coles, Globe Book Services, FREEPOST
Brunel Road, Houndmills
Basingstoke, Hants RG21 2BR England
Basingstoke (0256) 26222
FREEPOST applies to UK only

Postage & Packing US/Canada Added to invoice at cost, prepaid free UK/Rest of World £2 00 per order UK, £2 00 per book Continental Europe, £3 50 per book rest of world UK libraries free

Advertisement

PFIZER AWARDS 1986

Pfizer Central Research at Sandwich, Kent announced in October 1982 that they would make up to six new Awards annually to support young scientists who had carried out meritorious research at British Universities or equivalent Institutions The research should have potential application in the search for human or animal health drugs In 1986 each award is £2,000.

The award winners are selected by a panel of Pfizer scientists in consultation with senior academic advisers. They are made on the basis of work published in the previous two years or, in the case of a second consecutive award, in the previous twelve months. No person may receive more than two awards. The awards are for the personal use of the recipients to further their research and are NOT open to application

The 1986 winners are.

Chemistry

1. Professor P. J. Kocienski, Department of Chemistry, University of Southampton.

"For his novel synthetic routes to spiroketals employed in his syntheses of (±) talaromycin B and (+) milbemycin β3".

Professor Kocienski also won a Pfizer Award in 1983

2. Dr E. J. Thomas, Dyson Perrins Laboratory, University of Oxford

"For his studies on the intramolecular Diels-Alder approach to the "carbocyclic" cytochalasans, for example in his synthesis of (–) proxiphomin".

Biology

4. Dr Christine Beedham, School of Studies in Pharmaceutical Chemistry, University of Bradford.

"For her contribution to the understanding of the metabolism of aza–heterocyclic compounds by the molybdenum hydroxylase enzymes"

5. Dr. Clive Robinson, Department of Clinical Pharmacology, Southampton General Hospital, University of Southampton

"For his work on synthesis and metabolism of eicosanoids in lung and the putative role of inflammatory mediators in bronchial asthma".

Pharmaceutical Sciences

3 Dr. John N. Staniforth, School of Pharmacy and Pharmacology, University of Bath.

"For his work on the physical properties of particulate solids used in pharmaceutical formulation design".

Animal Health Research

6. Dr Q. A. McKellar, Department of Veterinary Pharmacology, Faculty of Veterinary Medicine, University of Glasgow

"For his work on ostertagiasis with particular reference to the interaction between Ostertagia ostertagi parasites and the gastric mucosa of cattle".

Pfizer is a worldwide, research-based company with inter-related businesses in health care, agricultural, specialty chemicals, materials science and consumer products. The Company has a major commitment at Sandwich where it employs approximately 1600 people, about 600 of whom are engaged in new product research.

AIDS in an African context

Clues to the origin of the virus that causes AIDS continue to emerge from studies of African viruses. But African countries have more pressing concerns

SIR Fred Hoyle may hold the view that the AIDS (acquired immune deficiency syndrome) virus is of extraterrestrial origin, others may like to believe that it escaped from a germ-warfare laboratory (surely a better plot would involve an unscrupulous biotechnology company that made and released the virus but whose vaccine development department has been a failure), but in all likelihood the human AIDS virus is a descendant of a monkey virus. The link between the two may be the second type of AIDS virus, identified earlier this year in western Africa. That is one, but by no means the only, reason why there is an increasing interest in the study of AIDS in Africa.

The reason to believe that the second type of virus is a link between monkey viruses and HIV-1 (human immune deficiency virus type 1, otherwise acronymically termed HTLV-3, LAV or ARV) is that it is more closely related to the former than the latter, at least by relatively crude comparisons. But the situation is complicated by the isolation of two apparently different versions of the west African AIDS virus and of several apparently different monkey viruses.

On the evidence so far published, the two versions of the human virus are very similar in composition but only one causes AIDS. That one was isolated by Luc Montagnier's group from AIDS patients in Guinea Bissau and the Cape Verde Islands, and was named LAV-2 by them (Clavel, F. *et al. Science* **233**, 343–346, 1986). The group has since also isolated it from AIDS patients in neighbouring countries, renamed it HIV-2 and, as reported further on in this issue (Clavel, F. *et al. Nature* **324**, 691–695, 1986), molecularly cloned the virus. The other version of the virus has been found in nearby Senegal by Myron Essex's group and always in people without AIDS (Kanki, P. J. *et al. Science* **232**, 238–243, 1986). This virus is called HTLV-4.

Unfortunately, and largely because of the sorry state of affairs that exists between leading French and US AIDS researchers, there has been no direct comparison of HIV-2 and HTLV-4. Therefore it remains uncertain whether their obvious similarities betoken an identity — or rather a pseudoidentity, because, as Clavel *et al.* now report, different isolates of HIV-2 vary in much the same way as those of HIV-1. The more interesting possibility is that HTLV-4 differs from HIV-2 in some crucial way that accounts for its apparent lack of pathogenicity.

Whatever the relationship of HTLV-4 and HIV-2, both seem more closely related to monkey viruses than they are to HIV-1. With their molecularly cloned HIV-2, Clavel *et al.* now confirm that relationship in terms of the degree of hybridization between the nucleic acids of HIV-2 and the virus that is the cause of the AIDS-like disease of captive rhesus macaque monkeys. This virus, usually known as STLV-3$_{\text{mac}}$ but now called SIV (unwisely without a species designation) by Clavel *et al.*, is related both to a virus in healthy captive sooty mangabey monkeys and to a virus in healthy wild African green monkeys. It is with the latter, called STLV-3$_{\text{agm}}$ and now molecularly cloned (Hirsch, V. *et al. Proc. natn. acad. Sci. U.S.A.* **83**, 9754–9758, 1986), that HTLV-4 has been compared and shown to be closely related. Hence the hypothesis that an apparently harmless monkey virus, STLV-3$_{\text{agm}}$, was the origin of an equally harmless human virus, HTLV-4, which evolved into the pathogenic HIV-2, the ancestor of HIV-1. Although this hypothesis is not without its problems, it also leads to some predictions that will be put to the test once all the viruses are cloned and sequenced. So far only HIV-1 is sequenced but the HIV-2 sequence is well on the way and the others should follow soon, if all goes well.

It is conceivable, at least, that this line of research will provide compelling evidence of an African origin of AIDS. Although that would not exactly be welcome news in Africa, at least the time is past when such a notion could lead to vehement denials, largely for fear of stigmatization and, it seems, loss of tourists. With a few exceptions, African countries are now frank about their AIDS problem, as well they might be given its scale.

Worst affected are central African countries, of which Kenya, Zaire and Uganda are the best studied. Reliable figures, which have been hard to come by, are summarized in two recent reviews (Quinn, T. C. *et al. Science* **234**, 955–963, 1986 and *AIDS and the Third World*, The Panos Institute*, 1986). As in developed nations, most of the data come from the major cities where the problem is at its greatest. The highest rates of infection are invariably in prostitutes, with figures ranging from 27 to 88 per cent. More alarming are the figures of 10–20 per cent from blood donors and antenatal clinics, as these are not high-risk groups in Western terms.

All the African figures have to be taken in context. Is, for example, the high rate of infection in blood donors typical of the population from which they are drawn or is it exaggerated by infection in the course of previous blood donations? Are the high rates of infection in general the result of other sexually transmitted diseases which, because of the genital ulcerations they cause, facilitate infection by the AIDS virus? And what rate of increase underlies these figures? Unfortunately there are few longitudinal studies of infection in Africa yet, although Quinn *et al.* boldly state that the present annual incidence of infection is approximately 0.75 per cent among the general population of central and east Africa. Nor are there many random population surveys, but a 6 per cent rate of infection recorded in a neighbourhood survey in Kinshasa, Zaire, is an indication of how serious the problem is.

More serious still is the fact that even the most simple means of limiting the spread of infection are not routinely used in most central African countries. It is estimated that between 1,000 and 1,500 new infections a year would be prevented in one major hospital alone in Zaire if there was routine screening of blood donors for HIV-1. Only a few million dollars would pay for routine blood screening throughout central Africa. But the cost per test would be somewhere between 3 and 30 times the average sum spent on an individual's health care in Africa. The prospects of routine testing seem poor without the help of foreign aid. To put the overall financial problem even more starkly, according to Quinn *et al.*, "the cost of caring for ten AIDS patients in the United States (approximately $450,000) is greater than the entire budget of a large hospital in Zaire, where up to 25 per cent of the pediatric and adult hospital admissions have HIV infection."

It is in this context that it is hardest to have much sympathy with some of the endless wranglings that go on concerning nomenclature of the viruses, with the neocolonial way in which some of the foreign interest in African AIDS is manifested, and with the lack of cooperation between rival groups when cooperation would further hasten understanding.

Peter Newmark

---

*8 Alfred Place, London WC1E 7EB UK, 1405 King Street, Alexandria, Virginia 22314, USA.

Fuel technology

Control of nitrogen oxides

A. Williams

AIR pollution arising from the emission of nitrogen oxides, NO_x, from combustion taking place in boilers, furnaces and engines, has increasingly been recognized as a problem (see, for example, *Fourth Report from the Environment Committee, Acid Rain*, HMSO, 1985). NO_x has been implicated as a photochemical oxidant and a constituent of acid rain, as well as being a pollutant in its own right. As a consequence, stringent legislation to limit its emission has been introduced or is proposed in the next few years in many parts of the world, but particularly in the United States, Japan and western Europe.

Several methods for controlling NO_x emissions have been developed (for example, *Control Technology for Nitrogen Oxide Emissions from Stationary Sources*, OECD, 1983). Most of these have some unsatisfactory aspects, however, and there is a constant search for new control technologies. Elsewhere in this issue R. A. Perry and D. L. Siebers (*Nature* **324**, 657, 1986) propose such a new technique. This method, developed at the Sandia National Laboratories, Livermore, California, is based on adding isocyanuric acid to hot exhaust gases, the acid being produced from the thermal decomposition of non-toxic cyanuric acid above 330 °C

$$(HOCN)_3 \rightarrow 3\,HNCO$$

Because this new control method has some attractive features it is worthwhile looking at it in the context of the control strategies that are currently available. The engineering aspects of the control strategies available for both NO_x and SO_2 are under investigation by the Fellowship of Engineering in the United Kingdom and its results will be published next year.

The major sources of NO_x are stationary combustion sources, such as boilers and furnaces, and motor vehicles, although some NO_x is emitted from chemical processing plants. The primary pollutant produced is NO, although some NO_2 and N_2O can be formed, and the NO reacts in the air to form a mixture of NO and NO_2 (NO_x). First, thermal NO is formed in combustion chambers from the oxidation of N_2 by the oxygen present in the 'excess air' needed to carry combustion to completion. Second, NO is formed from fuel-nitrogen compounds present in coal or to a smaller extent in oil, but not in natural gas. Third, NO can be formed in fuel-rich flames where hydrocarbon radicals react with nitrogen, this source, which is normally small, is termed 'prompt-NO'.

Various NO_x control technologies have been directed at combustion chamber modifications or a flue-gas treatment called denitrification. The combustion chamber modifications involve techniques to reduce the flame temperature and/or to reduce the amount of oxygen available for primary combustion. In engines this can be achieved by lean-burn techniques or stratified combustion chambers. In boilers and furnaces it can be achieved by low NO_x burners that depend on carefully controlled mixing of fuel and air, by flue-gas recirculation, or by low-temperature combustion techniques such as fluidized bed combustion or catalytic combustion.

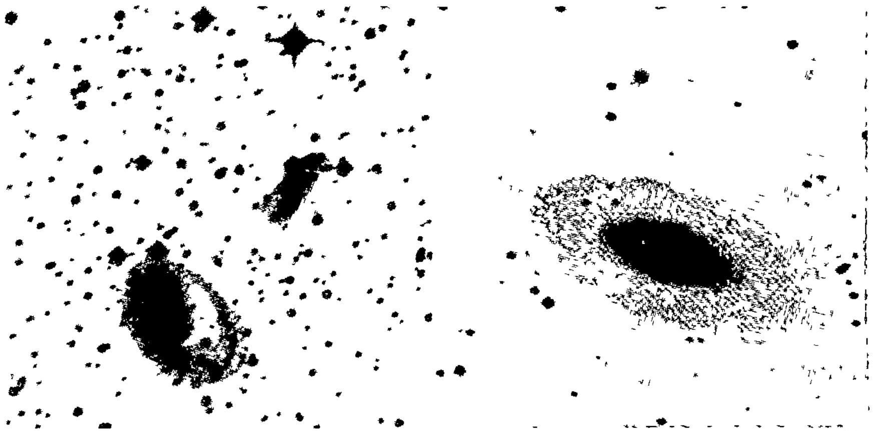
The NO_x regulations currently proposed are so strict that in the case of both engines and stationary combustion units it is becoming increasingly difficult to obtain the low concentrations of NO_x required by combustion modifications alone. Consequently, interest has centred on flue-gas denitrification for supplementing or replacing combustion control methods.

The methods available depend on the application and particularly whether the gases are clean or dust-laden. The most attractive methods involve (1) heterogeneous catalytic destruction usually based on precious metal catalysts and involving the reaction $2\,NO \rightarrow N_2+O_2$, (2) ammonia injection, called thermal de-NO_x, based mainly on the reaction $NH_2 + NO \rightarrow N_2 + H_2O$ and so on, (3) a catalytic version of the ammonia process usually called selective catalytic reduction, and (4) injecting methane or similar hydrocarbons to 'reburn' NO by reactions such as $CH + NO \rightarrow HCN \rightarrow CO + N_2$. The first of these methods needs a cleanish gas, such as engine exhaust or the flue gases from natural gas combustion, but needs controlled oxygen (air) addition. The other two need controlled ammonia or methane injection and controlled flue-gas temperatures and are really suitable only for large stationary plants. The applicability of many of these methods is limited where extensive heat recovery is required and the platinum-based catalysts, which operate down to 350 °C, look particularly attractive, if deactivation does not occur.

The new technique of Perry and Siebers therefore looks an attractive proposition. It provides a means of reducing NO_x levels down to those required by the most stringent regulations by passing the gas over the reactant without having to control the oxygen levels. It will operate over a temperature range of 450 to 900 °C, although at lower temperatures (100–450 °C) it actually produces NO_x itself. It seems clear that the work describes a technique that is a valuable additional NO_x control technology.

Their paper stems from earlier work (Perry, R. A. *J. chem. Phys.* **82**, 5485, 1985) that contains an interesting proposal about the chemistry leading to the formation of N_2O in some flames containing fuel-nitrogen. Although it has been known for several years that N_2O is formed during the combustion of organo-nitrites or nitrates (see, for example Gray, P. & Williams, A. *Nature* **188**, 56, 1960) it is a surprise that it can be formed in significant amounts as a result of the combustion of coal. The new work is therefore particularly interesting in that it proposes a control technology linked with these chemical routes and offers additional control mechanisms for future development. □

A. Williams is Professor of Fuel and Energy at the University of Leeds, Leeds LS2 9JT, UK.



JANE Few and Barry Madore, in a survey published in a recent issue of the *Monthly Notices of the Royal Astronomical Society* (222, 673–682; 1986), identify two sub-types of ring galaxies. On the left is a P-type ring galaxy showing the crisp, knotty structure of the ring itself, the displaced nucleus and companion galaxies. On the right is an O-type galaxy, showing the smooth structure of the ring and the centrally located nucleus. These data support the idea that P-type galaxies were formed from disk galaxies by the direct head-on collision with a companion and that O-type galaxies formed in isolation. □

Cellular signalling

A second messenger function for inositol tetrakisphosphate

Bob Michell

AN entirely novel concept entered our knowledge of how receptors send signals through membranes to the cell interior when it was realized that the receptor-activated hydrolysis of phosphatidylinositol 4,5-bisphosphate yields two second messengers rather than one Inositol 1,4,5-trisphosphate (Ins(1,4,5)P$_3$) and 1,2-diacylglycerol have different molecular targets within the cell, but they often act synergistically during the evocation of cell responses (reviewed in refs 1–3) Enzymic removal of the 5-phosphate of Ins(1,4,5)P$_3$, discovered in 1982, provided a perfectly respectable mechanism for its inactivation In the past year or so, however, it has become increasingly clear that there is a quantitatively important alternative pathway for the rapid metabolism of Ins(1,4,5)P$_3$ in stimulated cells, as a result of which it is phosphorylated to inositol 1,3,4,5-tetrakisphosphate (Ins(1,3,4,5)P$_4$), which is then dephosphorylated to inositol 1,3,4-trisphosphate, Ins(1,3,4)P$_3$ (see my recent News and Views article[4] and refs 5,6) It has been widely expected that Ins(1,3,4,5)P$_4$ or Ins(1,3,4)P$_3$, or maybe both, would function as additional intracellular messenger(s) generated by this increasingly complex signalling system, Irvine and Moor[5] now present the first evidence in support of this hope Using sea urchin eggs, they obtain evidence that strongly suggests that Ins(1,3,4,5)P$_4$ regulates the entry of calcium ions into these cells from the external medium

Whittaker and Irvine[7] were the first to show that injection of Ins(1,4,5)P$_3$ into sea urchin eggs suspended in a Ca^{2+}-containing medium causes the raising of a fertilization membrane, a response which seems identical to that produced by normal fertilization by sperm It is generally agreed that this response is mediated by a rise in intracellular Ca^{2+} that is fed both from internal Ca^{2+} stores and from the external medium Ins(1,3,4,5)P$_4$ does not mobilize Ca^{2+} from intracellular stores[6], so it was no surprise when it failed to activate eggs[5] But it has been clearly demonstrated that Ins(2,4,5)P$_3$ mobilizes Ca^{2+} from the same intracellular store as Ins(1,4,5)P$_3$, so its failure to raise a fertilization membrane was unexpected[5] Irvine and Moor surmise that the key to the difference between Ins(1,4,5)P$_3$ and Ins(2,4,5)P$_3$ might lie in the ability of the former to be converted to Ins(1,3,4,5)P$_4$ in the eggs into which it is injected, and wondered whether activation requires the simultaneous presence within the eggs of Ins(1,4,5)P$_3$ and of Ins(1,3,4,5)P$_4$ derived from it (Ins(2,4,5)P$_3$ is not an effective alternative substrate for Ins(1,4,5)P$_3$ 3-kinase, so it could not provide Ins(2,3,4,5)P$_4$ as an Ins(1,3,4,5)P$_4$ analogue) To test this idea, eggs were incubated in a Ca^{2+}-containing medium and injected with a mixture of Ins(1,3,4,5)P$_4$ and Ins(2,4,5)P$_3$ The results are startling This mixture of two ineffective inositol phosphates activates the eggs as effectively as Ins(1,4,5)P$_3$ alone As with other modes of egg activation, success with the Ins(1,3,4,5)P$_4$/Ins(2,4,5)P$_3$ mixture requires the presence of extracellular Ca^{2+} Irvine and Moor conclude[5] that Ins-(1,3,4,5)P$_4$ is likely to be the intracellular messenger responsible for initiating Ca^{2+} entry from the exterior, and that this is somehow dependent on the prior or simultaneous Ins(1,4,5)P$_3$-triggered discharge of the intracellular Ca^{2+} store They envisage that these two processes are coupled in some mandatory way such as that envisaged by Putney[8] his idea is that Ca^{2+} entering from the exterior does not move directly into the cytosol, but is routed through the Ins (1,4,5)P$_3$-sensitive pool, with influx particularly rapid when this pool is either partially or fully discharged

If future data confirm this view of events in sea urchin eggs and extend it to other cells, we now have all the key elements necessary to understand receptor-regulated Ca^{2+} mobilization into the cytosol of most stimulated cells Of course, Ins(1,4,5)P$_3$ and Ins(1,3,4,5)P$_4$ need to be both produced and inactivated in a coordinated way A remarkably tidy facet of this entire system is that the inositol polyphosphate 5-phosphatase that was originally detected as an Ins(1,4,5)P$_3$ 5-phosphatase may simultaneously inactivate both of these signals □

1 Berridge, M J & Irvine, R F *Nature* **312**, 315 – 321 (1984)
2 Nishizuka, Y *Nature* **308**, 693 – 698 (1984)
3 Downes, C P & Michell, R H in *Molecular Mechanisms of Transmembrane Signalling* (eds Cohen, P & Housley, M D) 3 - 56 (Elsevier, Amsterdam, 1985)
4 Michell, B *Nature* **319**, 176 – 177 (1986)
5 Irvine, R F & Moor, R M *Biochem J* **240**, 917 – 920 (1986)
6 Irvine, R F , Letcher, A J Lander, D J & Berridge, M J *Biochem J* **240**, 301 – 304 (1986)
7 Whittaker, M & Irvine, R F *Nature* **312**, 636 – 639 (1984)
8 Putney, J W Jr *Cell Calcium* 7, 1 – 12 (1986)

Bob Michell is in the Department of Biochemistry at the University of Birmingham, PO Box 363, Birmingham B15 2TT, UK

Atomic motion

Quantum fluctuations in sub-micron wires

Patrick A Lee

PHYSICISTS have long held the notion that measurements on a macroscopic sample are the same as the average over an ensemble of similarly prepared samples In the past two or three years, this notion has been challenged by a combination of two factors the ability to prepare small samples using micro-fabrication techniques, and the general availability of low temperature in the tens of millikelvin range The typical sample dimension is 0 05 µm diameter and 1 µm long, so that these 'small' samples still contain 10^8–10^9 atoms The surprise is that these samples, which are macroscopic on the atomic scale, already exhibit large fluctuations in their electric conductivity which also vary from sample to sample Furthermore, theoretical developments have led to a quantitative description of this phenomenon, and point to the possibility of detecting the change in the conductivity of a small metallic sample if the sample is perturbed in the slightest manner, even in the limit of displacing a single atom

These developments were initiated by the discovery that the resistances of small metallic wires (0 04 × 0 04 × 0 7 µm^3) exhibit random structure when a magnetic field is applied perpendicular to the wires[1] The structure is observable below 1 K, it increases in magnitude down to 10 mK where it is about a 0 1 per cent effect, it is reproducible upon cycling of temperature and magnetic fields, and it is a signature of a particular sample

At these low temperatures, the resistance is entirely due to the presence of impurities and imperfection in the sample and the motion of the electron through the sample must be described in terms of quantum mechanics A particularly illuminating way of describing the low temperature conductance G (inverse of resistance) of a disordered sample was provided by Landauer[2], who showed that G is proportional to the transmission probability of the electron wave through the disordered medium The proportionality constant is e^2/h which is a combination of fundamental constants with the dimension of conductance and equals 4 × 10^{-5} Ω^{-1} (The same combination is the unit of the Hall conductance in the quantized Hall effect)

The first step towards understanding the experiment was taken by Stone[3], who performed numerical calculations based on the Landauer formula for two-dimensional samples consisting of 40 × 400 sites and found fluctuations qualitatively similar to the experimental observation The idea is that according to Feynman's formulation of quantum mechanics, the transmission probability is given by a sum of the quantum amplitudes for many classical paths, and that the phase of each path is changed by the magnetic field in a random way As a result the transmission probability and therefore G exhibit fluctuations as a function of magnetic field However, numerical models showed a 10 per cent fluctuation and it was not clear how one could extrapolate results based on 10^4 sites to the experimental system with 10^8 atoms One might have guessed that as the sample size increases, the fluctuation will average away, but the numerical work showed that the magnitude of the fluctuation is almost independent of sample size

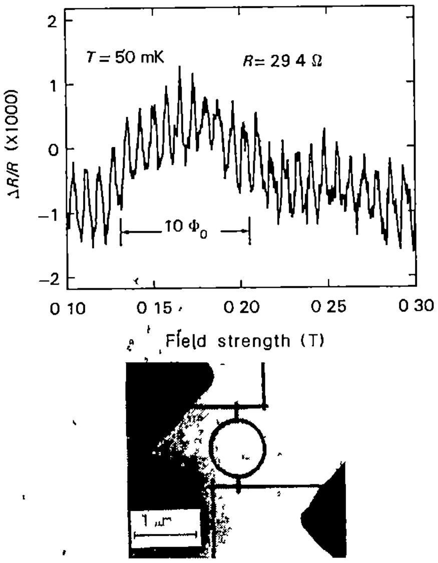


Fig 1 Variation of the resistance (R) of a metallic ring (photographed in the inset) with magnetic field The field strength is measured in tesla and is also shown in terms of the flux through the ring in units of the flux quantum $\Phi_0=h/e$ The oscillation with period h/e is the Aharonov–Bohm effect and the random structure it rides on is the universal conductance fluctuation The magnitude of both are of the order of $e^2/h \sim 4 \times 10^{-5}\ \Omega^{-1}$ in conductance units The inside diameter of the ring is 0 784 μm and the width of the wires is 0 041 μm (From ref 9)

The answer to these questions was provided by Lee and Stone[4] and independently by Altshuler[5] who showed analytically that if the question is posed in terms of ΔG, the root-mean-squre fluctuation in the conductance, instead of the percentage fluctuation,then at sufficiently low temperature ΔG is always of the order of e^2/h, independent of sample size and the amount of disorder, as long as disorder is weak and the sample is metallic The requirement of low temperature is that the electron should not suffer any inelastic collision as it traverses the sample For a 1-μm structure, this requirement is usually satisfied below 10 or 100 mK The prediction of this phenomenon of universal conductance fluctuation is consistent with both the numerical and experimental work and has now been verified in various experimental systems[6,7]

Universal conductance fluctuation is a consequence of quantum-mechanical coherence on a macroscopic length scale A most dramatic manifestation of this coherence is the Aharonov–Bohm[8] effect in normal metals, which was observed at about the same time as universal conductance fluctuation[9] (The Aharonov–Bohm effect is the idea that the electromagnetic potential alone, in the absence of any electromagnetic field, causes an observable quantum effect — a shift in the phase of the wavefunction) As shown in Fig 1, a small metallic ring is fabricated and its conductance is measured as a function of magnetic field perpendicular to the ring By extensions of Aharonov and Bohm's original arguments[8] which applied to electron motion in free space, it was predicted theoretically[10] that electron waves which propagate down the two arms of the ring would be phase-shifted by the magnetic field, and their interference would lead to periodic oscillation of the conductance The observation of these periodic oscillations (Fig 1) is a dramatic demonstration of quantum-mechanical coherence

One might worry that the amplitude of the periodic oscillations may be greatly reduced due to cancellations between the many Feynman paths with random phase which propagate down the arms But this is the same issue addressed by the theory of universal conductance fluctuation Thus, the magnitude of the periodic oscillation should be the same as the random structure, that is, of the order of e^2/h This is indeed the case for the data in Fig 1

The remarkable sensitivity of the conductance of a single sample to the magnetic field raises the following question how sensitive is the conductance to a change in the atomic configuration of a sample? The question was answered by Altshuler and Spivak[11] and independently by Feng, myself and Stone[12] who showed that changing the location of a single atom in a sample can, under suitable circumstances, lead to a conductance change of e^2/h as well and that this change is independent of sample area for thin films

This remarkable result is a natural consequence of the view of the electron motion as a superposition of many classical Feynman paths The point is that the classical motion of an electron across a strongly disordered metal is a random walk, which explores a large portion of the sample In fact, in two dimensions (thin films) the random walker visits a sizeable fraction of all the impurities, so that disturbing a single impurity affects the phase of a sizeable fraction of all the Feynman paths In other words, changing a single impurity has the same effect as replacing the entire sample According to the theory of universal conductance fluctuation, this leads to a change of the conductance of the order of e^2/h for highly disordered metallic thin films and wires Less dramatic, but still significant changes are predicted for bulk samples as well

Ideally, one would like to design an experiment to test this prediction, but nature appears to be kind enough to have provided an example Experimentalists at Purdue University[13], studying the resistivity of small wires and thin films, were perplexed by their observation that at temperatures below 1 K their data were not reproducible from run to run A more careful examination of the resistance as a function of time shows that it is time dependent, Figure 2 shows an example where the resistance makes two excursions in the period of an hour In this example, the resistance returns to the base line after a few minutes, but more complicated time dependence is also observed The experimentalists interpreted these jumps as being caused by the motion of a single atom or a small group of atoms and the magnitude of the jump is consistent with the theoretical prediction If this interpretation is correct, the motion of a single atom in a macroscopic metallic sample is indeed detectable It is not known what caused the atomic migration in the bismuth films at these low temperatures, but the effect may allow study of the dynamics of such slow configuration changes

Although the initial discovery of sample-specific fluctuations was made at low temperatures, it has become clear that quantum-mechanical coherence plays an important role as long as the inelastic scattering length greatly exceeds the mean free path, a condition which is satisfied

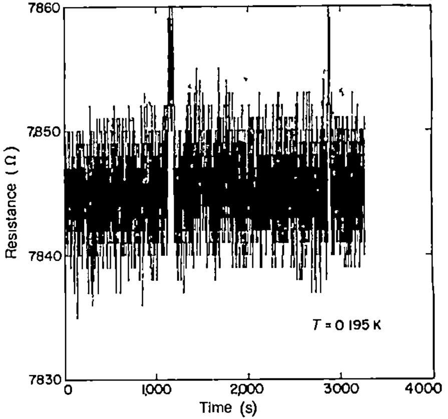


Fig 2 The resistance of a bismuth film (16 × 8 × 0 007 μm^3) at 0 195 K as a function of time in seconds The two excursions from the constant background at 1,100 and 2,900 s (rapid fluctuations are instrumental noise) are interpreted as being due to the motion of a single defect in the bismuth film (From ref 13)

even at room temperature if the metal is sufficiently disordered For example, Feng *et al* [12] point out that the sensitivity to impurity configuration provides an estimate of the magnitude of 1/*f* noise due to defect motion in highly disordered metals At the same time, it is likely that insights gained from the studies of electron wavefunctions will deepen our understanding of the propagation of classical waves such as electromagnetic or sound waves through a random medium □

Patrick A Lee is in the Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

1 Umbach, C P, Washburn, S, Laibowitz R B & Webb R A *Phys Rev* B30, 4048 (1984)
2 Landauer, R *IBM J Res Dev* **1**, 223 (1957)
3 Stone, A D *Phys Rev Lett* **54**, 2692 (1985)
4 Lee, P A & Stone, A D *Phys Rev Lett* **55**, 1622 (1985)
5 Altshuler, B L *Pis'ma Zh Eksp Teor Fiz* **41**, 530 (1985), *JETP Lett* **41**, 648 (1985)
6 Skocpol, W *et al Phys Rev Lett* **56** 2865 (1986)
7 Licini J C, Bishop, D Kastner, M A & Melngailis, J *Phys Rev Lett* **55**, 2987 (1985)
8 Aharonov, Y & Bohm, D *Phys Rev* **115**, 485 (1959)
9 Webb, R A *Phys Rev Lett* **54** 2696 (1985)
10 Buttiker, M, Imry, Y & Landauer, R *Phys Lett* A96, 365 (1983)
11 Altshuler, B L & Spivak, B Z *Pis'ma Zh Eksp Teor Fiz* **42**, 294 (1985), *JETP Lett* **41**, 363 (1985)
12 Feng, S, Lee, P A & Stone, A D *Phys Rev Lett* **56**, 1960 (1986), see also erratum *Phys Rev Lett* **56**, 2772 (1986)
13 Bentler, D E, Meisenheimer, T L & Giordano, N *Phys Rev Lett* (submitted)

Carcinogenesis

A superfamily of potentially oncogenic hormone receptors

Stephen Green and Pierre Chambon

UNDERSTANDING how patterns of gene transcription are controlled during the development of eukaryotic organisms and in terminally differentiated cells is an important goal of molecular biology This control appears to be achieved through the interaction of *trans*-acting proteins with *cis*-acting DNA promoter elements, which may result in a positive or a negative effect (see ref 1) Very few regulatory proteins have been characterized in higher eukaryotes Steroid hormone receptors are such a family of regulatory proteins, whose ability to control gene expression is dependent on the binding of their specific ligand Activation of transcription seems to result from the specific binding of the hormone–receptor complex to promoter 'enhancer' elements of target genes[2] Thus, the study of these receptors is of interest not only for elucidating how the receptor interacts specifically with its ligand, but also for understanding how this binding leads to the activation of transcription through specific protein–DNA interactions New work from four laboratories reported in this issue[3-6] brings us closer to understanding these mechanisms but also produces some apparent contradictions

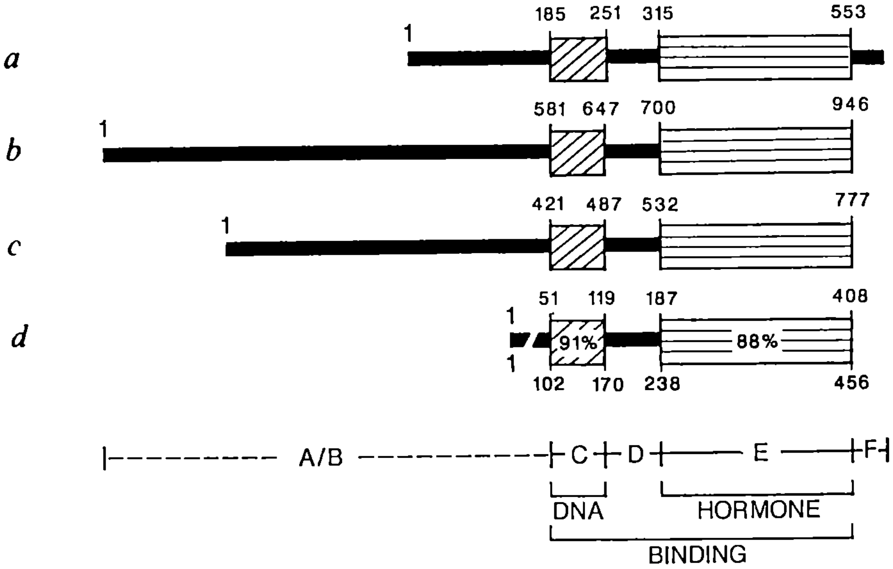
The recent cloning of glucocorticoid[7,8], oestrogen[9,10] and progesterone[11,12] receptors provides the first opportunity to investigate, at the molecular level, how transcription by RNA polymerase class B (II) is regulated in higher eukaryotes Sequence comparisons[10], together with mutation analysis[13,14], have defined two conserved regions that correspond to important functional domains (see figure) The putative DNA-binding domain is a highly conserved 66-amino-acid region (C) which has the potential to form two DNA-binding 'fingers' analogous to those proposed for the transcription factor TFIIIA (see refs 15 and 16), but involving two pairs of cysteines instead of pairs of cysteines and histidines[11] Results obtained with a chimaeric receptor show that region C determines the specificity of the receptor for the transcription of target genes[17] The hormone-binding domain (region E) is located in the C-terminal half of each of the receptors[13,14] The hydrophilic region D, which is variable in both length and sequence, may correspond to a hinge between the DNA- and hormone-binding domains Whether region A/B is required for activating transcription is not clear because analysis using insertional mutants of the human glucocorticoid receptor[14] indicates its importance for activation of transcription, whereas results obtained with deletion mutants of the rat glucocorticoid[18] and human oestrogen (Kumar, V *et al*, in preparation) receptors suggest that it is dispensable

Sequence comparisons revealed similarities between the product of the v-*erb*-A gene of the avian erythroblastosis virus and the human glucocorticoid[19] and oestrogen[9] receptors (see figure) This similarity is most striking in region C which is equally conserved when comparing these receptors and v-*erb*-A (about 50–60 per cent in each case) Because significant homology is also found in region E (although to a lesser extent, see figure), it was suggested that c-*erb*-A, the cellular counterpart of v-*erb*-A, is a receptor for a steroid-related ligand[10] Two reports on pages 635 and 641 of this issue[3,4] demonstrate, surprisingly, that the product of both a chicken[3] and a human[4] c-*erb*-A gene bind specifically thyroid hormones with high affinity (although the affinity with which the two receptors bind T_3 differs by an order of magnitude), and may therefore be the chicken and the human receptors for those hormones This possibility is further supported by the nuclear localization of the chicken c-*erb*-A product

The thyroid hormone receptor is known to be localized in the nucleus and it is believed that it modulates gene expression by a mechanism similar to that of steroid hormone receptors (see ref 20) Regions C and E of the thyroid hormone receptor may thus correspond to their DNA- and hormone-binding domains, respectively As region C of either the human and chicken oestrogen receptors[10] or the human and rat glucocorticoid receptors[21] are entirely conserved, it is surprising that C regions of the chicken c-*erb*-A and the human c-*erb*-A are only 91 per cent homologous, a value which is identical to the homology found between the progesterone and glucocorticoid receptors (see figure) These two c-*erb*-A products may thus correspond to different genes encoding receptors that bind the same hormones, but exhibit different target gene specificites This possibility is supported by the finding of several *erb*-A-related genes located on difficult chromosomes in the human genome[4] Studies on the transcriptional activation of thyroid hormone receptor-responsive genes using the two cloned c-*erb*-As are required to demonstrate definitively that they correspond to functional thyroid hormone receptors That steroid and thyroid hormones, which are neither structurally nor biosynthetically related, have receptors that have evolved from a common ancestor gene, suggests strongly that there is a large superfamily of genes whose products are transcriptional regulatory proteins All these proteins probably have similar DNA-binding domains, but it will not be surprising to find that the activity of some of them may have become independent of ligand binding, thus allowing them to act as constitutive transcription factors This superfamily could be comparable with that of the homoeobox-containing proteins that are also involved in the control of gene expression during development, but whose putative DNA-binding domain may contain the prokaryotic helix-turn-helix DNA-binding motif rather than DNA-binding fingers[16]

The classical model for steroid hormone action proposes that the hormone is required for translocation of the receptors to the nucleus and for its binding to the DNA-responsive elements of target genes (see ref 2) This model is based on the observation that the hormone is required for the receptor to bind tightly to the nucleus However, it has been shown more recently that the oestrogen (see ref 22) and the progesterone (see ref 23) receptors are localized, but weakly bound, in the nucleus even in the absence of hormone

What, then, is the role of the hormone? Could the hormone-free receptor, for instance, occupy the DNA steroid-



Schematic alignment of steroid and thyroid hormone receptor sequences The amino-acid sequences of the human oestrogen receptor (*a*), chicken progesterone receptor (*b*), human glucocorticoid receptor (*c*), and both the chicken (1–408) and human (1–456) c-*erb*-A proteins (*d*), were deduced from their complementary DNA sequences Alignment is based on amino-acid similarity and the numbers refer to the position of amino-acid residues The division of the steroid hormone receptors into 6 regions (A–F), based on a comparison between the chicken and human oestrogen receptors[10], is shown at the bottom of the figure The two highly conserved regions, representing the putative DNA-binding (region C) and hormone-binding (region E) domains, are shown as shaded blocks Little or no significant homology is observed when comparing other regions of the receptors (solid black lines) Comparison between the chicken and human c-*erb*-A protein sequences shows 91 and 88 per cent sequence identity in regions C and E, respectively Comparison between the various receptors in regions C and E show the following homologies (in per cent) C, steroid receptors/c-*erb*-A ~ 45–55, human oestrogen/chicken progesterone or human glucocorticoid receptors, 62, chicken progesterone/human glucocorticoid receptors, 91 Region E steroid receptors/c-*erb*-A ~20, human oestrogen/chicken progesterone or human glucocorticoid receptors, ~30, chicken progesterone/human glucocorticoid receptors, 55

responsive elements of target genes without activating their transcription? This important question is addressed in two papers on pages 686 and 688 of this issue[5,6] and in another which has just appeared in the *EMBO Journal*[23] Using the *in vivo* dimethylsulphate footprinting technique, the group of Schutz[5] shows that occupancy of the glucocorticoid-responsive element of the glucocorticoid-inducible liver tyrosine aminotransferase gene is influenced by the presence of the hormone and of its antagonist RU486 Protection of guanine residues within the glucocorticoid-responsive element is observed only in the presence of dexamethasone and not in the absence of hormone nor in the presence of RU486 Although a weak interaction of hormone-free glucocorticoid receptor with the responsive element cannot be excluded by these experiments, the results clearly support the notion that dexamethasone increases the affinity of the receptor for its target sequence

In contrast, the two other reports present *in vitro* data which suggest that glucocorticoid[6] and progesterone[23] receptors recognize preferentially their specific DNA-binding sites, irrespective of the presence of the hormone, and even when complexed with their common antagonist RU486 The group of Milgrom[24], which has studied the binding of highly purified hormone-free rabbit progesterone receptor to specific fragments of the target uteroglobin gene, concludes that the affinity of the receptor for its target sequence is not modified by the presence of the hormone or of its antagonist, and that specific recognition is achieved even at 0°C Willmann and Beato[6] have approached the same problem by studying the ability of crude rat liver cytosolic preparations of hormone-free glucocorticoid receptor to bind to the responsive element of the mouse mammary tumour virus long terminal repeat sequence Again, no effect of the hormone, or its antagonist, on the DNA-binding properties of the cytosolic receptor is found, although a heat-activation step (25 °C, 30 min) of the cytosol preparation is essential Thus, it appears that the hormone is not necessary for specific DNA-binding *in vitro*, whereas it is *in vivo*

There are at least two explanations that can reconcile these conflicting data As proposed by both Willmann and Beato[6] and Bailly *et al*[23], the receptors may be complexed *in vivo* with some protein(s) or structure(s), preventing their binding to the target-gene responsive elements Binding of the hormone to the receptor would induce the dissociation of such a complex, allowing the receptor to interact with its target sequences One possible candidate for such an inhibitor is the heat-shock protein of relative molecular mass 90,000 which, under certain conditions, readily associates with several steroid receptors[24] Alternatively, it cannot be excluded that extensive purification of the receptor or an *in vitro* heat-activation step may mimic the action of the hormone by artefactually inducing a conformational modification of the receptor required for recognition of specific target sequences This modification could correspond either to the 'unmasking' of a previously hidden, but potentially active, DNA-binding domain, or to a conformational activation of a previously inactive domain However, true affinity measurements (based on the 'on' and 'off' rates of the receptor to its target sequence) in the presence and absence of the hormone have not been performed in either of the two *in vitro* studies[6,23] This may explain some discrepancies between the results obtained *in vitro* by Willmann and Beato[6] and those previously reported by Bourgeois *et al*[25], who have also studied the effect of dexamethasone and RU486 on glucocorticoid-receptor binding to the mouse mammary tumour virus long terminal repeat Thus, until real affinity measurements have been performed, it still remains an open question as to whether the efficiency of receptor binding to responsive elements is independent of the presence of the hormone *in vitro* In fact the recent studies of Beato's group (personal communication) indicate that the *in vitro* 'on' and 'off' rates may be higher in the presence of the hormone This may result in a marked effect on the occupancy of responsive elements by the receptors *in vivo* where the number of non-specific, compared with specific, DNA-binding sites is very large

The avian v-*erb*-A oncogene, which cannot by itself induce cell transformation, enhances the erythroblast transforming potential of several primary oncogenes, such as v-*erb*-B (see refs 26 and 27) The v-*erb*-A gene product completely inhibits erythroblast differentiation into erythrocytes and allows transformed erythroblasts to grow in standard tissue-culture media, which has led to the notion[3] that in erythroid cells, v-*erb*-A interferes with ion transport systems and regulatory factors of differentiation Because of the sequence similarity between v-*erb*-A and steroid hormone receptors, it has been speculated[10,19,28] that v-*erb*-A exerts its effect either by mimicking or by competitively antagonizing the effect of c-*erb*-A on the transcription of target genes Assuming that the thyroid and steroid hormone receptors function similarly, it is intriguing that v-*erb*-A can exert its specific effects on erythroblasts even though it cannot bind thyroid hormones, presumably because of some of the many mutations present in its putative ligand-binding domain (region E)[3] It is likely that, because of other mutations[27], v-*erb*-A has acquired the ability to bind 'constitutively' to responsive elements of target genes Whether this would result in v-*erb*-A mimicking or competitively inhibiting (for example, because the truncation of region A/B in v-*erb*-A would prevent it from activating transcription) the effect of c-*erb*-A, and/or whether the two amino-acid substitutions that are present in the putative DNA-binding domain of v-*erb*-A[3] have resulted in changing its target gene speicificity when compared with c-*erb*-A, remains to be seen

If one assumes that the thyroid hormone receptor acts like steroid hormone receptors, the reports of Weinberger *et al*[4] and Sap *et al*[3] in this issue suggest that 'altered' enhancer factors may be important in oncogenic transformation by interfering with the transcriptional regulation of crucial target genes The well-known effect of the presence of oestrogen receptor on the growth of some human breast cancers (see ref 29) and the insertion of the hepatitis B viral genome into an *erb*-A-related gene of a human hepatoma[30] suggests that the superfamily of receptor genes contains several potential oncogenes encoding enhancer binding factors The polyoma virus enhancer can be activated *in trans* by expression of the c-Ha-*ras* oncogene and by the tumour promoter TPA[31] Furthermore, the activity of several viral and cellular enhancers can be repressed *in trans* by expression of the adenovirus immortalizing oncogene *E1A* (see ref 32) It may be possible, therefore, that alteration of enhancer elements of key cellular genes and/or of their binding factors correspond to crucial steps in many types of carcinogenesis □

1 Ptashne, M *Nature* **322** 697 – 701 (1986)
2 Yamamoto, K R *A Rev Genet* **19**, 209 – 252 (1985)
3 Sap, J *et al* *Nature* **324**, 635 – 640 (1986)
4 Weinberger, C *et al* *Nature* **324**, 641 – 646 (1986)
5 Becker, P B , Gloss, B , Schmid, W , Strahle, U & Schütz, G *Nature* **324**, 686 – 688 (1986)
6 Willmann, T & Beato, M *Nature* **324**, 688 – 691 (1986)
7 Miesfeld, R *et al* *Nature* **312**, 779 – 781 (1984)
8 Hollenberg, S M *et al* *Nature* **318**, 635 – 641 (1985)
9 Green, S *et al* *Nature* **320**, 134 – 139 (1986)
10 Krust, A *et al* *EMBO J* **5**, 891 – 897 (1986)
11 Jeltsch, J M *et al* *Proc natn Acad Sci U S A* **83**, 5424 – 5428 (1986)
12 Conneely, O M *et al* *Science* **233**, 767 – 770 (1986)
13 Kumar, V , Green, S , Staub, A & Chambon, P *EMBO J* **5**, 2231 – 2236 (1986)
14 Giguère, V , Hollenberg, S M , Rosenfeld, M G & Evans, R M *Cell* **46**, 645 – 652 (1986)
15 Miller, J , McLachlan, A D & Klug, A *EMBO J* **4**, 1609 – 1614 (1985)
16 Berg, J M *Nature* **319**, 264 – 265 (1986)
17 Green, S & Chambon, P *Nature* (in the press)
18 Rusconi, S *et al* in *RNA Polymerase and the Regulation of Transcription* (eds Reznikoff, W S *et al*) (Elsevier, New York, in the press)
19 Weinberger, C , Hollenberg, S M , Rosenfeld, M G & Evans, R M *Nature* **318**, 670 – 672 (1985)
20 Oppenheimer, J H & Samuels, H H (eds) *Molecular Basis of Thyroid Hormone Action* (Academic, New York, 1983)
21 Miesfeld, R *et al* *Cell* **46**, 389 – 399 (1986)
22 Schrader, W T *Nature* **308**, 17 – 18 (1984)
23 Bailly, A , Le Page, C , Rauch, M & Milgrom, E *EMBO J* **5**, 3235–3241 (1986)
24 Jaob, I *et al* *Nature* **308**, 850 – 853 (1984)
25 Bourgeois, S , Pfahl, M & Baulieu, E E *EMBO J* **3**, 751 – 755 (1984)
26 Kahn, P *et al* *Cell* **45**, 349 – 356 (1986)
27 Damm, K , Beng, H , Graf, T & Vennström, B *EMBO J* (in the press)
28 Bishop, J M *Nature* **321**, 112 – 113 (1986)
29 Edwards, D P , Chamness, G C & McGuire, W L *Biochim biophys Acta* **560**, 457 – 486 (1979)
30, DeJean, A , Bougueleret, L , Gzeschick, K H & Tiollais, P *Nature* **322**, 70 – 72 (1986)
31 Wasylyk, C , Imler, J L , Perez-Motul, J & Wasylyk, B *Cell* (in the press)
32 Hen, R , Borrelli, E & Chambon, P *Science* **320**, 1391 – 1394 (1985)

Stephen Green and Pierre Chambon are in the Laboratoire de Génétique Molécular des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'Inserm, Institut de Chimie Biologique, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cédex, France

Cosmology

Views with a gravitational lens

Rachel Webster and Michael Fitchett

GRAVITATIONAL lensing should in principle provide a fundamentally new technique for exploring the distribution of mass in the Universe, for determining cosmological parameters and for determining the properties of background sources For example, model fitting can determine the mass distributions of specific lenses and statistical studies can be used to determine the existence and density of populations of compact objects, strings or clumped dark matter In theory, constraints can be placed on the cosmological parameters H_0, Ω and Λ, but the realization of such measures may be difficult The amplification bias introduced by gravitational lensing will affect source counts of background objects and may either be skewing what we observe, or may serve as a natural telescope to probe the structure of objects at high redshift

The central point that emerged from a recent workshop*, however, was the need for verification that supposed gravitational lenses are indeed what they have been claimed to be Because one can in principle learn a great deal from either a particular instance of gravitational lensing, or a well-defined sample of multiply imaged quasars, it is important that a strict procedure is developed to confirm the status of a lens When the first candidate lens, 0957+561A,B, was discovered[1] in 1979, it was subjected to rigorous scrutiny Observations in a range of wavelengths and the identification of a reasonable candidate for the lens object have served to strengthen the case that the two images observed are of the same quasar Of course, if one hopes to learn something about the distribution of dark matter, the identification of a candidate lens object should not be a requirement, and the emphasis switched to other properties of the system A further eight candidates have now been added to the list It was not until the wide separation lens 1146 + 111B,C (at an image separation of 157 arc s it is about 20 times larger than any of the previous candidates) was mooted[2] that rigorous testing was again considered

Both necessary and sufficient conditions for lens identification were discussed by Jacqueline Hewitt (Haystack Observatory) and Irwin Shapiro (Harvard) Multiple images are seen along different photon paths This gives rise to spectral differences between the images, a time delay and perhaps a velocity difference If there is intrinsic variation in the source, not only will the spectra be different at one instant in time, but each will also vary with time Variation in the lens potential will change the relative amplifications of the images with time, and in particular amplifications of different regions of the image Once the tolerances in velocity, spectra and spectral changes for two images of the same quasar have been calculated, a sufficient set of conditions for the positive identification of a lens needs to be determined Possible tests include redshift, spectra, image morphology, identification of the lensing object and its effect on other background sources, time delay, polarization and parity The last two tests can easily be applied only to radio quasars Where such data are available, as in the case of 0957+561A,B, the case for lensing becomes very strong For optical quasars however, if no deflecting object is seen, definitive identification relies not only on determination of the redshift and spectra, but also on the measurement of a time delay As there are ten known optical quasars to every one known radio quasar, this ratio will be reflected in gravitational lenses, and an efficient technique for verifying optical candidates is required

Reports of systematic surveys for multiple images by four groups were highlights of the meeting Hewitt reported on a survey[3] with the Very Large Array This collaboration has identified a good candidate lens, 2016 + 112, and Hewitt described a further two candidates This survey is a sensitive probe of matter clumped on scales of $10^{10} - 10^{12}$ times the mass of the Sun, and already the authors can derive a conservative upper limit of $\Omega_{Lens} \leqslant 0\ 7$ for masses in this range Marc Gorenstein (Harvard) described a new survey using archive Very Long Baseline Interferometry data which will be sensitive to multiple images in the range 0 1 – 1 arc s This is the range of separations expected by galaxies acting as lenses John Stocke (Joint Inst of Lab Astrophysics) reported a study of X-ray active galactic nuclei selected near bright foreground galaxies These sources show some evidence, both in redshift distribution and in the ratios of X-ray to optical fluxes, of being micro-lensed by stars in the foreground galaxy One of us (R W) described an automated survey for optical candidates for gravitational lensing This study uses the automated plate measuring machine at Cambridge, United Kingdom automatically to select quasar candidates with possible image multiplicity Already this study can place limits on the numbers of lenses with separations in the range 10 arc s to 2 arc min (there are no candidates in a sample of about 2,500 quasars), and it can accumulate a substantial statistical sample of gravitational lenses in time

---

*The gravitational lensing workshop was held at the Department of Astronomy, University of Toronto and the Canadian Institute for Theoretical Astrophysics on 3 – 4 October 1986

Apart from the surveys, it became apparent that two other systematic monitoring projects are required First, the known candidates need to be regularly monitored to determine time delays C Vanderreist (Meudon Observatory) presented data for 0957+561 collected over a six-year time base So far a unique time delay cannot be determined from the data

Second, as Shapiro emphasized, none of the data collected so far gives a definitive measure of the time delay, to make use of information about the individual variations more needs to be known about the general population of quasars Not only do we need to know how quasars vary spectrally one from another, but also we need to know more about the power spectrum of the intrinsic variability

Peter Schneider (Joint Inst of Lab Astrophysics) distinguished between linear and non-linear cases of lensing In the former case, the optical depth in lenses is small, and reasonable progress can be made in analytically solving problems such as the effect of gravitational lensing on source counts In the non-linear case more than one strong lens event affects the photon path, and global problems need to be solved numerically A particular example of the latter is micro-lensing where the photon beam passes through a dense field of compact objects A second area of theoretical work is the modelling of specific instances of multiple images In most cases the modelling proceeds by trial and error until a single solution is found Such solutions are never unique, and in fact no systematic method for examining the ensemble of posible solutions has been developed Roger Blandford (California Institute of Technology) and colleagues are evaluating image parameters for different lens potentials to allow the most probable mass distribution in the lensing object to be determined

Third, the effects of a clumped matter distribution on the measurements of cosmological parameters was discussed by Charles Dyer (Toronto) In particular, many weak lenses can significantly affect background source amplifications[4] Myeong-Gu Park (Princeton) described a calculation which uses the known separations of double images to place limits on the cosmological parameters Λ and q_0, in particular $q_0 \geqslant -2\ 35$

Four new applications of lensing were discussed Bohdan Paczyński (Princeton) discussed the possibility that some of the gamma-ray bursts observed may be multiple images of the same event Such an observation would be evidence for clustered micro-lenses, the main difficulty in verifying this hypothesis is the determination of an accurate position for each individual burst Robert Wagoner (Stanford) explored the possibility of using supernovae as probes of dark compact objects He presented detailed calculations of the light curves expected for Types I and II supernovae which had been micro-lensed, and dicussed the possibility of actually observing such an event in a distant galaxy Ulf Borgeest (Hamburg) described a parallax method, requiring a baseline of about 0 1 AU, which would allow the determination of both the size of the quasar emitting region as well as the transverse velocity of the lens Borgeest also reported a method for using time delays to determine the mass of the primary galaxy in the lens

The wide separations of the observed doubles, and the absence in most cases of a suitable candidate for the lensing object, pose significant problems Blandford pointed out that if quasars are clustered like galaxies then one might predict a few correlated quasar pairs in the known sample of quasars[5] If, however, the multiple images without suitable lens objects are gravitationally lensed, then the mass distributions of the lenses have high mass-to-light ratios and cannot be composed of normal stellar material Thus the verification of candidate lenses becomes of central importance ☐

1 Walsh, D *et al* *Nature* **279**, 381 (1979)
2 Turner, E L *et al* *Nature* **321**, 142 (1986)
3 Lawrence, C R *et al* *Science* **223**, 46 (1984)
4 Dyer, C C *Proc Vat Conf Theor Observational Cosmology* (in the press)
5 Phinney, E S & Blandford, R D *Nature* **321**, 569 (1986)

Rachel Webster is in the Astronomy Department and Michael Fitchett is in the Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario M5S 1A1, Canada

---

Developmental biology

New data from mammalian homoeobox-containing genes

Michael H L Snow

SERENDIPITY gave us the homoeobox[1], a highly conserved DNA sequence of about 180 base pairs, first found in several *Drosophila* genes that control segmentation Homoeoboxes were also detected in other segmented invertebrates and in several vertebrates including mammals, giving rise to the view[2] that this small DNA segment would serve to identify segmentation genes This notion has since been revised because echinoderms and molluscs, not generally regarded as segmented, also have homoeobox-containing genes[3] The recent demonstration[4] that the *Drosophila* homoeobox-containing gene *Zerknullt* (*zen*) is transcribed only in dorsal regions without segment specificity suggests that these genes have a wider scope than the organization of the anterior–posterior axis of the embryo The new data of Gaunt *et al* and Mavilo *et al* on pages 662 and 664 of this issue[5,6] add to our understanding of the specificity of expression of mouse[5] and human[6] homoeobox-containing genes

The possibility that homoeobox-containing genes specify positional information in the embryo is still high and it is this aspect of patterning in morphogenesis that underlies the considerable interest in

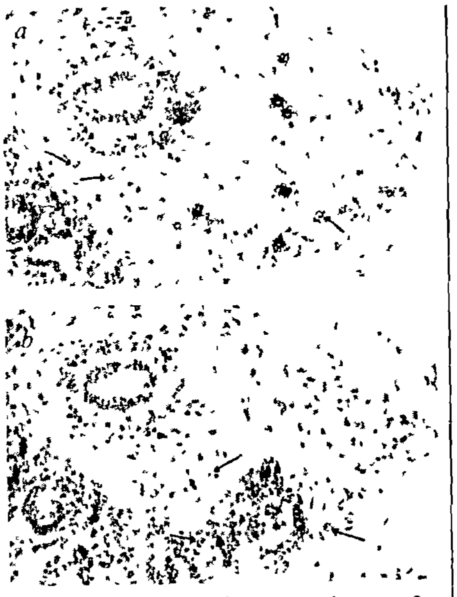
Mouse homoeobox-containing genes

Name of gene	Chromosome	Synonym	Ref	Best homologies (per cent nucleotides amino acids)		
				Drosophila	Human	*Xenopus*
Hox-1	6	*m6*	13	Antp(82 95)	Hu 2(80 88)	MM3(79 83)
		m5	14	Antp(81 95)	Hu 2(83 90)	MM3(82 93)
		m2	14	Antp(?? ??)	?	?
		Hox-1 3	15	Antp(73 82)	Hu 1(82 90)	MM3(74 82)
		Hox-1 4 (mo10)	2,5	Antp(70 74)	?	?
Hox-2	11	*Hox-2 1 (mu1, H24 1)*	16,17,18	Antp(87 90)	Hu 1(92 95)	MM1(? 87)
		Hox-2 2	16	Antp(80 ??)	Hu 2(91 100)	?
		Hox-2 3	16	Antp(77 ??)	?	?
		Hox-2 4	16	Antp(?? ??)	?	?
Hox-3	16	*Hox-3 (m31)*	9	Antp(73 80)	?	?
En-1	10	*Mo-en 1*	7	*en* (65 73)	?	?
En-2	5?	*Mo-en 2*	7	*en* (?? ??)	?	?

mammalian homoeoboxes. It has not been possible so far to analyse developmental mutants in mammals (principally in the mouse) with the thoroughness that has been applied to *Drosophila* but it is hoped that the homoeobox will enable us to identify and isolate such genes. So, how far have we got?

Six homoeobox-containing genes have been identified in the human and a dozen in the mouse (see table). As in *Drosophila*, these genes seem to be clustered and to show temporal and spatial restriction in transcription. *Hox-2.1* transcription, for example, is detectable in mouse embryos as early as 7.5 days post coitum (during late primitive streak stages of gastrulation), shows peak levels in 11.5–12.5-day-old embryos and is enriched in spinal cord. Recent work with *Hox-2.1* using Northern blot analysis and *in situ* hybridization detects transcripts in embryonic kidney and in the mesodermal component of the developing lung (P. Holland, B. Hogan and R. Krumlauf, in preparation) and show also that transcription is temporally related to the differentiation of F9 embryonal carcinoma (EC) cells (R. Krumlauf, personal communication). In this latter system the duration of the transcription phase and its position in the differentiation sequence is related to whether differentiation is towards visceral or towards parietal endoderm. The mouse *Mo-en.1* gene is also differentially expressed during differentiation of EC cells[7]. Gaunt *et al.*[5] in this issue describe similar differential expression of the mouse *Hox-1.4* (*Mo10*) gene by *in situ* hybridization. Here, transcription is again detectable at 7.5 days and is later found confined to spinal cord, mesonephric kidney and somites. The human homoeobox genes described by Mavilio *et al.*[6] in this issue likewise show considerable stage and organ specificity in expression. All mammalian homoeobox genes show similar restrictions in expression to some extent. What do the data tell us?

The short answer is very little beyond what I have discussed above. One problem is the difficulty in defining comparable aspects of mammalian and *Drosophila* development, and this is severely complicated by technical limitations in the analysis of the mammalian embryo. The extremely powerful genetic analyses in *Drosophila* support and are reinforced by the ability to detect the earliest transcriptional activity by Northern blot analysis, to locate it accurately by *in situ* hybridization, and to locate translation by detection of protein products. The data are consistent, showing the onset of gene activity — with known developmental consequences — at or shortly after cellularization of the blastoderm.



Kidney of 14.5-day-old mouse embryo. *a*, *In situ* hybridization with the anti-sense probe to mouse ε-globin. Heavy label localized over nucleated erythrocytes of yolk-sac origin. Arrows indicate three enucleated erythrocytes that originated from the liver. *b*, Control hybridization on adjacent section with the sense strand. Arrows indicate three nucleated yolk sac-derived erythrocytes which show up as positive in *a*. Probe donated by Andy McMahon, hybridization by Peter Koopman and Alex Reid.

The luxury of a four-pronged attack is not possible in the mammalian embryo. No genes are clearly definable as homoeotic or segmentation, although there are candidates for the latter[8], and no antisera against homoeobox gene products have yet been described. So we have only the relatively insensitive transcription analyses (Northern blots and RNase protection assays) and the potentially very useful *in situ* hybridization techniques. At the moment neither technique alone tells us very much. Transcription analysis requires that tissues can be dissected in sufficient quantity and is therefore applicable only to late embryonic or fetal stages if spatial distribution of transcripts is to be studied. This is long past the time that positional cues associated with lineage definition are likely to be needed. Furthermore, any small population of positive cells, for example, stem cells, may well be swamped by a mass of non-transcribing tissue. For earlier, more interesting stages, *in situ* hybridization is the technique of choice: the 7.5-day-old mouse embryo will yield less than 20 ng of messenger RNA and although it is possible to accumulate material for transcription analysis from parts of these embryos it is clearly a daunting task and not suited to the present impatient scramble to publish homoeobox data. The *in situ* hybridization technique is difficult to apply to mammalian embryonic tissue. Non-specific binding is a common, but recognizable, artefact[9] so it is imperative that rigorous control hybridizations, which should include a probe whose distribution is predictable in addition to the non-hybridizing, non-complementary 'sense' strand of the probe under study, are not only carried out but also displayed alongside the hybridization with the complementary anti-sense strand of interest.

It is possible, when cell discrimination is helped by histological parameters, to locate accurately *in situ* hybridization signals to individual cells (see figure), but where morphology does not assist it can be far from easy. Even if such data were to be acquired for 7.5-day-old mouse embryos it would then leave the not inconsiderable problem of relating any patterns found to known organ-specific transcription shown in fetal stages. The fate maps and lineage data that exist for these early stages only relate to development up to about the 10-somite stage[10], that is at least 24 hours before the primordia of many organs currently associated with homoeobox transcripts are found and 3–4 days earlier than the regional transcription analyses. Meanwhile, there are groups of apparently disparate organs expressing homoeobox genes at various developmental stages and we have little idea of what, if any, common feature they reflect. There are representatives of induced and inducing tissues, derivatives of different primary germ layers and organs which seem temporally and spatially unrelated in their origins. It may be a misleading coincidence, but the homoeobox-positive tissues equate roughly with those organs with determinate numbers of structural units, and negative tissues seem to be those that retain hyperplastic abilities throughout life[11].

Drosophila shows determinate development from the blastoderm stage. Although the information from *Drosophila* is not entirely straightforward[12], with curious gene interactions requiring explanation, the general picture is clear. The distribution of homoeobox gene transcripts is adequately linked through fate maps and genetics to later phenotype. There is some way to go before a similar state is reached with mammals and before we understand what their homoeobox-containing genes may do. ☐

1. Gehring, W.J. *Cell* **40**, 3–5 (1985).
2. McGinnis, W. *Cell* **37**, 403–408 (1984).
3. Holland, P.W.H. & Hogan, B.L.M. *Nature* **321**, 251–253 (1986).
4. Doyle, H.J. *et al. Nature* **323**, 76–79 (1986).
5. Gaunt, S.J. *et al. Nature* **324**, 662–664 (1986).
6. Mavilio, F. *et al. Nature* **324**, 664–668 (1986).
7. Joyner, A.L. *et al. Cell* **43**, 29–37 (1985).
8. Hogan, B. *et al. Trends Genet.* **1**, 67–74 (1985).
9. Awgulewitsch, A. *et al. Nature* **320**, 328–335 (1986).
10. Snow, M.H.L. *J. Embryol. exp. Morph.* **65** (Suppl.), 269–287 (1981).
11. Goss, R.J. *Science* **153**, 1615–1620 (1966).
12. Coulter, D. & Wieschaus, E. *Nature* **321**, 472–474 (1986).
13. Colberg-Poley, A.M. *et al. Nature* **314**, 713–718 (1985).
14. Colberg-Poley, A.M. *et al. Cell* **43**, 39–45 (1985).
15. Duboule, C.P. *et al. EMBO J.* **5**, 1229–1235 (1986).
16. Hart, C.P. *et al. Cell* **43**, 9–18 (1985).
17. Hauser, C.A. *et al. Cell* **43**, 19–28 (1985).
18. Jackson, I.J. *et al. Nature* **323**, 76–79 (1986).

Michael H. L. Snow is in the MRC Mammalian Development Unit, University College London, 4 Stephenson Way, London NW1 2HE, UK.

Mitochondrial DNA

Rearranging the plant genome

Ronald A. Butow

THE tendency of many plant mitochondrial genomes to undergo extensive rearrangements provides a satisfying framework for understanding the phenomenon of cytoplasmic male sterility (CMS), a maternally inherited defect common to many higher plants that impairs the development of pollen. Some extraordinary examples of the molecular basis for CMS in plants such as maize, sorghum and petunia were provided at a recent conference*.

The mitochondrial genomes of higher plants tend to be very large (200–2,400 kilobases) compared with other organisms. The presence of many direct repeats within some plant mitochondrial DNAs can lead to a profusion of subgenomic units that arise by recombinations across the repeat elements, as was shown for maize (D. Lonsdale, Plant Breeding Institute, Cambridge), wheat (F. Quetier, Université de Paris-Sud, L. Bonen, University of Ottawa) and *Oenothera* (A. Brennicke, Universitat Tubingen). A somewhat simpler picture obtains for species of *Brassica* the mitochondrial genomes of related species are roughly the same size, but they vary in the number and distance between repeats on the 'master' mitochondrial genome. This variation neatly accounts for the number, size and, in some cases, compete absence of subgenomic elements (J. Palmer, University of Michigan).

As if the presence of repeats leading to recombinational partitioning of plant mitochondrial genomes is not complex enough, on closer scrutiny, other sub-stoichiometric sets of differently rearranged species are also apparent in maize mitochondria (P. Issac, University of Edinburgh). These sub-stoichiometric populations, called sublimons, are maintained through many generations, and thus must somehow compete with the 'major' mitochondrial DNAs in that plant. Conventional thinking about the drive to segregate pure populations of mitochondrial genomes must be refined to consider the origin of sublimons by continued recombination or replicative persistence, and perhaps occasional amplification, in which they may eventually take over the population.

So where does CMS fit into this picture of a promiscuously rearranging genome? In maize, CMS is associated with C, T and S cytotypes, whereas normal male-fertile strains are designated N. Compared with N strains, mitochondrial DNA from C male-sterile cytoplasms show three striking rearrangements that result in the formation of chimaeric genes (R. Dewey, North Carolina State University) first, the 5′ flanking region of the ATPase 9 gene has been replaced by sequences of unknown origin, second, the 5′ ATPase 9 sequences plus a small amount of the coding region are now found upstream of an open reading frame (ORF) originating from the chloroplast genome, which, in turn, is fused in-frame to the ATPase 6 gene, and third, the replaced ATPase 6 sequences are fused in-frame to the cytochrome oxidase 6 gene, whereas the displaced cytochrome oxidase 6 sequences end up in a region with no ORF.

The mitochondrial genome from T strains which are not only male sterile, but are also susceptible to toxin from the fungal pathogen *Helminthosporium maydis*, contains a rearrangement that is perhaps more bizarre than C. No less than seven recombination events must have taken place to create a chimaeric gene containing ATPase 6 sequences, a region 3′ to 26S ribosomal DNA, a portion of the interior of 26S ribosomal DNA and a transfer RNA (Arg) derived from the chloroplast genome. Earlier work (C. Leaver, University of Edinburgh) showed that a novel polypeptide of relative molecular mass 13,000 was synthesized in the mitochondria of T strains. It now seems clear that this polypeptide is the product of this chimaeric gene, also called ORF 13. A precise correlation can now be made between restoration of fertility and expression of ORF 13. In one case of a fertile, toxin-resistant mutant of T, two changes were detected in ORF 13 a single G-to-A transition in the reading frame, and an adjacent 5-base pair insertion creating a frameshift, which may have occurred by a gene conversion from flanking DNA sequences 3′ to the 26S ribosomal RNA gene (R. P. Wise, University of Florida). In other cases, mutation to fertility is caused by additional rearrangements that result in the deletion of the novel gene (W. Rottman, Plant Breeding Institute, Cambridge, C. Fauron, University of Utah). Restoration of fertility can also come from post-transcriptional processing activities that reduce the levels of translatable ORF 13 messenger RNA (Dewey, A. Abbott, Clemson University).

The picture is less clear for CMS-S. Male steriles of the S type have two linear DNA plasmids in their mitochondria, S1 and S2, that can integrate into the mitochondrial genome. In some nuclear backgrounds, reversion to fertility is always associated with the loss of S1 and S2 (E. Earle, Cornell University), although this is not always the case in other nuclear backgrounds (S. Gabay-Laughnan, University of Illinois). S-specific gene products have been identified (G. Zabala, University of Illinois, Leaver), but where S1 and S2 are present in fertile revertants, there are no apparent differences in the synthesis of the plasmid-encoded proteins examined so far.

That the nuclear context can be important in expression of the CMS trait is also evident from studies with sorghum. A variant cytochrome oxidase subunit I gene has been identified (J. Bailey-Serres, University of California, Berkeley) containing both 5′ and 3′ rearrangements. These result in the synthesis of a long form of the protein (relative molecular mass 70,300) — the largest cytochrome oxidase subunit I polypeptide identified to date. Although the precise relationship of this rearrangement to CMS is not completely known, some lines containing the variant gene are male sterile whereas others are fertile. In either case, there are no obvious effects on cytochrome oxidase activity.

Petunia provides another example of the association of CMS with mitochondrial gene rearrangements (S. Izhar, Volcani Center, Israel, M. Hansen, J. Rasmussen and H. Nivison, Cornell). Here, recombination leads to the formation of a chimaeric gene between ATPase 9, cytochrome oxidase subunit II (COII) and DNA sequences from an unidentified reading frame. This fusion gene is expressed in male-sterile lines, as are the intact copies of ATPase 9 and COII genes. Expression of the chimaera seems to be unaffected in the presence of a dominant nuclear restorer gene, but preliminary evidence suggests normal COII expression is increased in a restorer background (Nivison).

It is clear from most of the evidence presented that some restorer genes function either by lowering or eliminating the expression of novel proteins encoded by rearranged mitochondrial DNA in male-sterile lines or, in some way, by overriding their presence. It seems reasonable to imagine that these novel proteins, which are probably located in the inner mitochondrial membrane, cause some mitochondrial dysfunction that becomes limiting at a critical step in microsporogenesis. But what this dysfunction might be is a complete mystery. In a related maternally inherited phenomenon in maize called non-chromosomal stripe, novel mitochondrial DNA rearrangements were identified — although not in any known genes — that result in severe alterations in mitochondrial (and chloroplast) morphology (K. Newton, University of Missouri). Because these plants show abnormal growth, a more immediate effect of mitochondrial dysfunction seems likely. ☐

Ronald A. Butow is in the Department of Biochemistry, University of Texas Health Science Center at Dallas, Texas 75235, USA

**International Workshop on Higher Plant Mitochondrial DNA* 19 – 24 October 1986, Airlie, Virginia, USA

Microtubule dynamics

SIR—Although Williams, Caplow and McIntosh[1] correctly emphasize the important contribution of Horio and Hotani[2] in verifying, by real-time experiments; the concept of the dynamic instability of microtubules, some of their points are off the mark and they have missed an opportunity to evaluate the proposed explanations for dynamic instability in terms of these and other recent experiments

First, dynamic instability was never meant to imply that shrinking microtubules would always depolymerize to completion Although we pointed out that the rate constant for this transition would be hard to measure, our own data suggested that microtubules could in fact transit from shrinking to growing[3,4] What would determine whether a microtubule disappears, in addition to chance, would be its length, the experimental conditions and, probably most importantly, the purity of the tubulin In the theoretical work of Chen and Hill recapping rates are also explicitly included[5] The unusual characterisitic of microtubules is that the growing and shrinking transients make such large excursions It is less surprising that in some circumstances microtubule number actually decreases Therefore there are no real inconsistencies between the Horio and Hotani data and ours, except perhaps that under their conditions the excursions are somewhat smaller than those indirectly estimated from our steady-state experiments

A second issue is whether length redistributions at steady state[4] are due to growing and shrinking transitions in microtubules or due to annealing of polymers Although Williams *et al* point out that the Horio and Hotani results dispel any doubt that extensive growing and shrinking transitions are continually occurring, they suggest that annealing could have contributed to the magnitude of our original findings They fail to mention, however, that in a subsequent paper annealing was found to make only a very small contribution to steady-state length redistributions[6] Annealing is an interesting reaction but most likely to occur far from steady state and at very high concentrations of microtubule ends

It is worth reconsidering the original GTP cap model[7] as it was applied to dyna/ic instability[4] in light of the new Horio and Hotani data One reason we favoured the GTP cap model was that, in order not to violate thermodynamic principles, energy input from GTP hydrolysis is required to support continuous growing and shrinking[8] A second reason was to explain our observation that the probability of a microtubule starting to shrink was inversely related to its growth rate In the Chen and Hill simulation[5] virtually all transitions must occur on the minus end, as it is the slow growing end Yet Horio and Hotani find transitions on both ends and even a greater frequency on the faster growing end This is inconsistent with the original model[9] where GTP hydrolysis occurs at a fixed rate once a subunit is incorporated into the lattice In this simple model (Fig 1*a*) a subunit in the interior of the polymer cannot know which end it is near and hence the hydrolysis rate of all subunits in the interior of the polymer would be the same The faster-growing end would leave behind it a larger number of GTP-containing subunits and the GTP cap should be lost more frequently on the slower-growing end Thus the slower-growing end necessarily has a higher probability of transition to shrinking behaviour

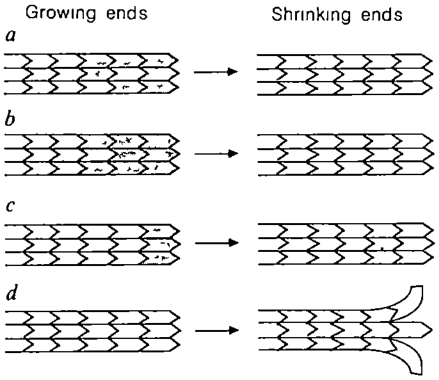


Fig. 1 Models for the molecular events that cause a growing microtubule end to start shrinking spontaneously The chevrons show the inherently asymmetric tubulin subunits in the lattice The blank chevrons are GDP-containing subunits and the stippled chevrons are GTP-containing subunits Note that the barbed end of the chevrons will be exposed at the other end of the microtubule (not shown)

Any explanation of Horio and Hotani's data must make use of structural differences at the two ends of the microtubule Presently we can see three ways of achieving this The first restricts GTP hydrolysis to the zone of subunits in contact with the GDP liganded core of the microtubule[1,10] In contrast to the previous model a subunit can tell which end it is near by which portion of the molecule is contacting subunits that have already hydrolysed their GTP In Fig 1*b*, for example, the pointed end of the GDP subunit is interacting with the barbed end of the GTP subunit, on the other end of the microtubule this relation is reversed This means that on one end of the microtubule (shown) GTP hydrolysis is induced by the pointed end of a GDP subunit, whereas on the other end (not shown), it is induced by the barbed end These different structural interactions could induce different rates of hydrolysis at the two ends, and thus the relation between growth rate and transition probability can be different at the two ends

The second possibility (Fig 1*c*) invokes a very minimal GTP cap, it is a model patterned after the conclusions of some recent work on actin by T Pollard[11] In his model, whether the polymer grows or shrinks depends only on the nucleoside composition of the terminal subunit or subunits Hydrolysis is effectively instantaneous for interior subunits, but slower for terminal subunits As the two ends of a polymer are structurally different, this could lead to different hydrolysis rates for terminal subunits at the two ends, and again to a different relationship between transition probabilities and growth rates at the two ends The complexity of microtubule structure means that there are many more possible terminal configurations than for actin filaments and this might extend the concentration range of free tubulin over which transitions would be observed Furthermore, rapid shrinking might generate conformations that make recapping difficult

The third possibility uncouples GTP hydrolysis from a structural transition at the end of the polymer lattice which initiates rapid disassembly (Fig 1*d*) The terminal subunits retain the GTP conformation for a period after hydrolysis (which is concomitant with assembly) and shrinking behaviour is initiated by a spontaneous fluctuation whose rate could be different at the two ends of the polymer The real situation could of course include elements from more than one model

One of the unique features of all these models for dynamic instability is that the stability of the polymer is regulated by reactions at the very ends of the microtubules But this potentially important biological property[8] is also an experimental difficulty as at any given time almost all subunits in the polymer and in solution are not carrying out important reactions The work of Horio and Hotani has dramatically demonstrated that the behaviour of individual polymers cannot be inferred from the behaviour of bulk populations It has further focused attention on the structure and enzymatic properties of the ends of polymers, a problem whose solution will require yet more ingenious experimentation

MARC W KIRSCHNER
TIM MITCHISON
*Department of Biochemistry
and Biophysics,
University of California
School of Medicine, San Francisco,
California 94143-0448, USA*

1 Williams, R J Jr, Caplow, M & McIntosh, J R *Nature* **324**, 106–107 (1986)
2 Horio, T & Hotani, H *Nature* **321**, 605–607 (1986)
3 Mitchison, T & Kirschner, M *Nature* **312**, 232–237 (1984)
4 Mitchison, T & Kirschner, M *Nature* **312**, 232–242 (1984)
5 Chen, Y -D & Hill, T L *Proc natn Acad Sci U S A* **82**, 4127–4131 (1985)
6 Kristofferson, D *et al* *J Cell Biol* **102**, 1007–1019 (1986)
7 Hill, T L & Carlier, M -F *Proc natn Acad Sci, U S A* **80**, 7234–7238 (1983)
8 Kirschner, M & Mitchison, T *Cell* **45**, 329–342 (1986)
9 Carlier, M -F & Pantaloni, D *Biochemistry* **20**, 1918 (1981)
10 Pantaloni D , *et al* *Proc natn Acad Sci U S A* **82**, 7207-7211 (1985)
11 Pollard, T *J Cell Biol* (in the press)

Preservation in bogs

SIR—Doran *et al* (*Nature* **323**, 803, 1986) have shown reasonable histological preservation of 8,000-year-old brain tissue from a bog site, yet they have not given an explanation for why the tissue preservation should be so good We suggest that it is due to tannic acid and related compounds contained in the bog fluids We have demonstrated excellent ultrastructural preservation of osteocytes from the 11,000-year-old antlers of an extinct Giant Irish Deer In this case preservation in the bog environment is so good that cytological features, such as mitochondria, rough endoplasmic reticulum, Golgi elements and nuclear pores are clearly demonstrated

GLENN R DICKSON
E MARTYN JOPE
Medical Biology Centre,
The Queen's University of Belfast,
Belfast BT9 7BL, UK
ALAN BOYDE
Department of Anatomy and Embryology
University College London,
Gower Street,
London WC1E 6BT, UK

How many reactor accidents?

SIR—Edwards[1] is correct that Islam and Lindgren[2] have failed to recognize that their distribution function is a likelihood function But he is incorrect in asserting that their conclusions are unwarranted and in assuming that physicists do not know how to work with small samples In fact Islam and Lindgren's likelihood function has been used in particle physics for over fifty years to make estimates of particle decay rates when the number of events in an experiment is small[3] Their approach yields not only estimates of the accident rate and probabilities but also the familiar standard errors and confidence limits without resort to Edwards' less familiar support function

The estimate of r, the accident rate, is given by the maximum of the likelihood function The standard error is conventionally quoted as the value(s) of r for which the likelihood L, falls to $e^{-1/2}$ of its maximum value And the (one-tailed) confidence limits on r are set by solving for r in the integral equation,

$$\int_r^{\infty} L_N \int \mathrm{d}r = P$$

where L_N is the normalized likelihood function These values of r can then be used to obtain estimates, standard errors and confidence limits on the probability P_T of accidents within the next T years

The two events over a cumulative 4,279 reactor years of operation[4] give the estimate $r=1/2{,}140^{+2\ 620}_{-1\ 000}$ reactor years with a one tailed 95% confidence limit $r>1/5{,}240$ reactor years In the next ten years 28 reactors are expected to be decommissioned and 156 are expected to come on line giving 4,900 reactor years of operation[4] The probability of one or more accidents in the next ten years is then $P_{10}=0\ 90^{+0\ 09}_{-0\ 26}$ with a 95% confidence level of $P_{10}>0\ 60$ Thus one has a quantitative estimate of how small the probability of accident is likely to be

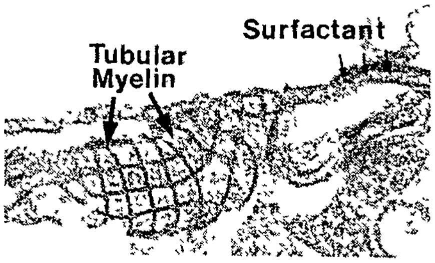
Given that there have been two major accidents in the world's reactor inventory over 4,280 reactor years of operation the odds are 20 1 ($p>95\%$) that in 100 identical ensembles of reactors 60 ensembles will have one or more accidents within ten years A similar calculation shows that even if the world's reactors are phased out over the next 25 years the probability of major accident is $P_{25}=0\ 994^{+0\ 005}_{-0\ 09}$ with $P_{25}>0\ 88$ at 95% confidence

JOSEPH SCHWARTZ
2 Lancaster Drive,
London NW3 4HA, UK

1 Edwards, A W F *Nature* **324**, 417–418 (1986)
2 Islam, S & Lindgren, K *Nature* **322**, 691–692 (1986)
3 Solmitz, F T A *Rev nuc Sci* **14**, 375–402 (1965)
4 *Nucl Engng Int (UK)* **30**, 7–15 (1985)

Alternative solutions

SIR—The small grid-like mystery object of Wolstenholme *et al*[1] may be a fragment of 'tubular myelin', which is the contracted or non-spread form of lipoproteinaceous pulmonary surfactant Fetal surfactant has been identified in human amniotic fluid[2], and the tubular myelin form is commonly observed The dark-stained squares probably represent protein which is perhaps stained during chromosome preparation



Transmission electron micrograph of tubular myelin in the lung of a rat exposed to asbestos Taken from ref 3

ARNOLD R BRODY
LILA H OVERBY
Laboratory of Pulmonary Pathobiology,
National Institute of Environmental
Health Sciences, PO Box 12233,
Research Triangle Park,
North Carolina 27709, USA

1 Wolstenholme, J *et al* *Nature* **323**, 300 (1986)
2 Hook, G E *et al* *Am Rev resp Dis* **117**, 541-550 (1978)
3 Brody, A R *et al* *Fedn Proc* **44**, 2596-2601 (1985)

SIR—Wolstenholme *et al*[1] have probably observed a well-preserved area of a meshwork of the nuclear lamina lining the nucleoplasmic surface of the inner nuclear membrane which supports the interphase chromatin It is unusual to observe any remnant of lamina in a metaphase preparation, but two chromosomes in the figure appear to be sticking to lamina

Full boxes in the pattern could represent the cytogenetic manifestation equivalent to the points of anchorage of the interphase chromosomes (telomeric ends and fragile sites) onto the laminar surface, thus reflecting the packing of chromosomes during interphase in a 'coded' form The average crossover space of the human lamina is 0 52 µm A 'decoded' picture of the pattern of native nuclear lamina of *Xenopus laevis* oocytes can be found in a past issue of this journal[2]

G M JANKOVIĆ
Clinic for Haematology,
Faculty of Medicine,
11 000 Belgrade,
ul Koče Todorovića bb, Yugoslavia

1 Wolstenholme, J *et al* *Nature* **323**, 300 (1986)
2 Aebi, U *et al* *Nature* **323**, 560 (1986)

Archaeopteryx, the primordial bird?

SIR—In view of the highly critical review of our book *Archaeopteryx, The Primordial Bird A Case of Fossil Forgery* (*Nature* **324**, 185, 1986) I would like to bring to your notice recent developments that would seem to have vindicated our views

The issue hinges on whether or not the material on which feather lines are impressed, and of which a few extraneous blobs exist elsewhere on the fossil surface, is comprised of a cement which is foreign to the rock At our request the museum authorities supplied us with a sample of the suspect material and of the native rock matrix These samples have now been examined by Dr Lee M Spetner and his colleagues in Israel using a scanning electron microscope and X-ray spectroscopy The sample from the rock matrix shows a characteristic crystalline structure exactly as in other specimens of Solnhofen limestone, with identical X-ray resonance spectra But the sample from the suspect material shows a non-cystalline structure resembling that of amorphous material bound by an organic glue X-ray resonance spectra showed large amounts of silicon as well as lead and chlorine which are certainly alien to native Solnhofen limestone

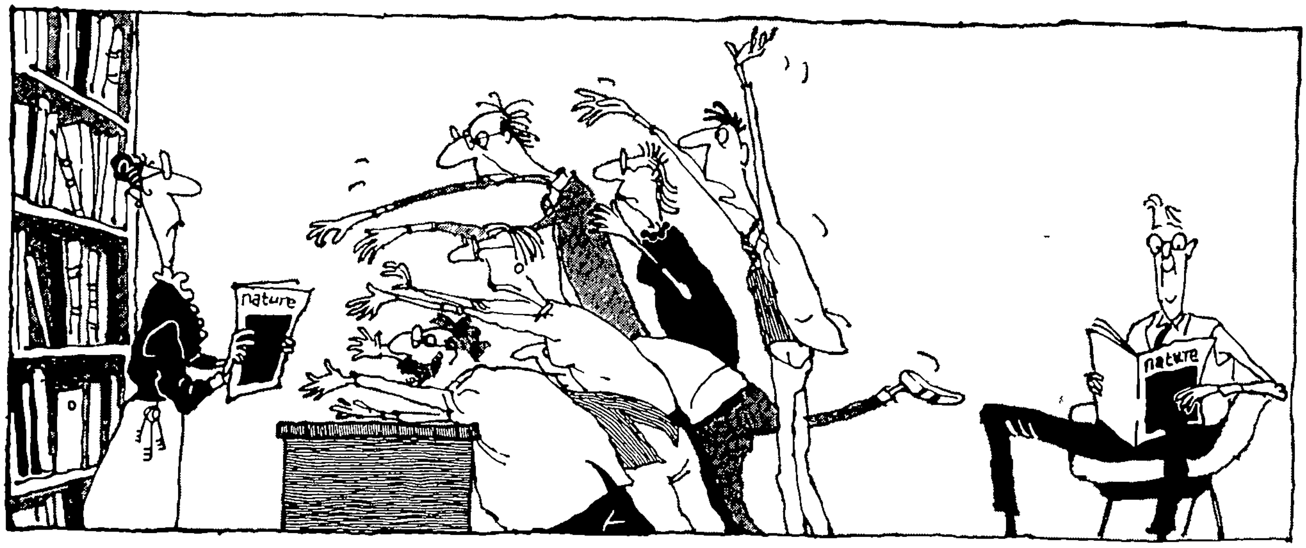
These striking differences in texture and composition between the suspect regions and the native matrix are, in our view, a strong indication that this dispute will eventually be resolved in our favour The results have been communicated to the British Museum along with a request for further samples of material from the fossil for examination with a view to narrowing down the possible organic compounds that are involved in the 'glue'

N WICKRAMASINGHE
F HOYLE
University College,
Cardiff CF1 1XL, UK

Why YOU should become a personal subscriber to Nature in 1987

There will never be a better opportunity than now to take advantage of our personal subscription scheme..... offering direct delivery to your home or to your place of work at HALF PRICE!

Consider the convenience of having your own personal copy delivered each week carrying crucial research results, news and job opportunities in your specific field HOT OFF THE PRESS! When **Nature** is such a vital ingredient in your scientific career a subscription is a very small investment to make – in fact the cost works out at less than your daily newspaper



All you need to do to subscribe is fill in the form below and send it with your personal payment FREEPOST to our offices in Basingstoke You'll receive your first copy within four to six weeks of your order and thereafter weekly for a year We *guarantee* that when we invite you to renew your subscription it will be at the lowest rate available

Order Form

Return to Circulation Manager, FREEPOST Macmillan Journals Ltd, Houndmills, Basingstoke, Hants, RG21 2XS

☐ YES, please enter my personal subscription to **Nature** at £52

☐ I enclose my cheque for £52 made payable to **Nature**

☐ Please charge my credit card account

☐ Access ☐ American Express

☐ Diners Club ☐ Barclaycard

Account No ____________ Expiry Date ____________

Signature ______________________

Name ______________________

Address ______________________

____

____

Please allow 4–6 weeks for delivery of first issue This offer applies to UK residents only – overseas subscription rates available from above address

7B66

This year treat yourself to something different. . . .

OCEANUS, the only magazine that seriously covers the three quarters of the world that most other magazines neglect

"For anyone who loves the sea and whatever moves upon it, OCEANUS seems to me a 'must' " WALTER SULLIVAN, The New York Times

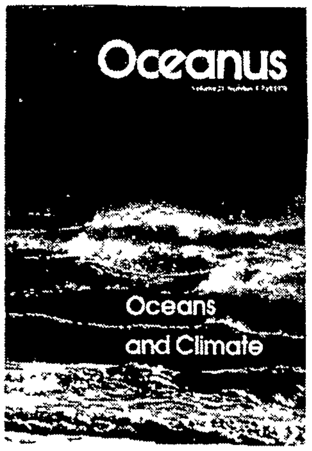
"OCEANUS makes it possible for the public to keep up with the most important work in marine science – and with the social implications of that work I recommend it highly" JEROME WIESNER, Former President, MIT

" the editors live in the same buildings as most of the scientists and the result is readable and attractive good science " THE CO-EVOLUTION QUARTERLY

OCEANUS covers marine science and technology, oil and gas activity, deep ocean mining, fisheries development, coastal zone management, conservation, ocean economies, marine dumping Get first-hand reports as the Woods Hole submersible **Alvin** gathers evidence for the theories of sea-floor spreading, turns up new life forms, explores the Titanic

Each exciting article is illustrated with some of the best photographs and art available

Subscribe from
U.S.A. – Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543
Rest of the World – (£20 for individuals, £37 for institutions) Journals Dept, Cambridge University Press, The Edinburgh Building, Shaftesbury Road, Cambridge, CB2 2RU, England



Reader Service No 3

JOURNALS FROM MACMILLAN

NEW JOURNAL FIRST ISSUE PUBLISHED OCTOBER 1985

☐ **ANTI-CANCER DRUG DESIGN**
Quarterly
Editors S Neidle S T Crooke

NEW JOURNAL FIRST ISSUE MAY 1986

☐ **BONE MARROW TRANSPLANTATION**
Quarterly
Editor J M Goldman
Co-Editor R P Gale

NEW JOURNAL LAUNCHED DECEMBER 1984

☐ **PHARMACEUTICAL MEDICINE**
Editor R N Smith
Assistant Editor P J Keen

☐ **BRITISH JOURNAL OF CANCER** *Monthly*
The Journal of the Cancer Research Campaign
Chairman D G Harnden Editor M Moore

☐ **BRITISH JOURNAL OF PHARMACOLOGY**
Monthly
The Journal of the British Pharmacological Society
Chairman J F Mitchell
Secretary to the Editorial Board G M Lees

☐ **PUBLIC HEALTH** *Bi-monthly*
Journal of the Society of Community Medicine
Editor J S Robertson

☐ **POSTGRADUATE MEDICAL JOURNAL**
Monthly
The Journal of the Fellowship of Postgraduate Medicine
Editor B I Hoffbrand

☐ **HUMAN TOXICOLOGY** *Bi-Monthly*
Secretary to the Editorial Board P Turner

For sample copies and details of subscription and agency rates, please tick the titles you are interested in, complete the details below and return to Macmillan Press at the address shown

Name ______________________________

Position ______________________________

Address ______________________________

____

Return to **Alison Baverstock,**
Macmillan Press, Houndmills,
Basingstoke, Hants,
RG21 2XS, UK

Macmillan Press
Scientific & Medical

Who shall be the judge?

Frederick Dainton

The Science Critic: A Critical Analysis of the Popular Presentation of Science.
By Maurice Goldsmith *Routledge & Kegan Paul 1986 Pp 217 £15 95, $35*

It is a truism, forever fixed in our childhood minds by the story of Pandora's box, that knowledge can be used for good or ill and that the problem for human societies is to maximize the former and minimize the latter In the eighteenth century this problem was sufficiently evident to Dr Johnson, when he astounded the world by completing his book *Rasselas* within a week, to invent a philosopher who would not communicate his secret of manned flight until mankind became more virtuous

Since 1759 there have been many other authors who have espoused the argument that scientists ought to cease work and suppress their findings as soon as it is clear that damage to their fellow men might ensue I doubt if these arguments would have carried much weight in Dr Johnson's day, in a society that was largely rural and agricultural and for which the relatively feeble power of the human biped was augmented only by that of large mammals or that available from wind or falling water Johnson's younger contemporaries such as James Watt, Abraham Darby and Matthew Boulton changed this situation irreversibly by applying knowledge, largely empirically gained, to invent the steam engine and iron and steel rails and wheels The revolution they started gathered momentum over about one and a half centuries and transformed Western society into a primarily industrial and urban one

Amongst the by-products of the Industrial Revolution were Victorian self-confidence and the establishment of colleges and institutions where science, engineering and technology could be taught and advanced in the sure and certain hope that the knowledge would further mankind's power over the animate and inanimate world, then deemed to be the only route to human material progress But there was a darker side which soon became evident and was manifested by environmental pollution and consequent disease, social strain and sometimes discord, and struggles between imperialist powers for access to the cheap raw materials and the captive markets necessary to maintain growth in the manufacturing industries It is arguable that neither the Industrial Revolution nor its outcomes would have been foreseen either by those pioneers who made it possible or the nation's political leaders

Now the society created by the Industrial Revolution is giving way to another, and the engine of this change is science The principal advances in knowledge are twofold first, the powerful combination of the analysis of operations into a series of questions demanding a simple yes or no answer with the developments in solid state physics to make the micro-chip possible, and, secondly, the understanding of the control and information systems in living cells The former has led to the computer, which has greatly augmented human mental power, whilst the latter has far-reaching implications for medicine, agriculture and the efficient and cheap production of chemicals It seems likely that the judgement of history will be that this Information Revolution will be much more significant than the Industrial Revolution It certainly poses weightier questions, for the potential benefits are so much larger and so are the risks

In the twentieth century scientists have shown themselves to be aware of the dilemma which this knowledge poses for them Some, like the founding members of the American Association of Atomic Scientists 40 years or so ago, thought they must make a concerted effort to explain to the public what the implications of unlocking the power of the atomic nucleus were In so doing they were adopting the view, later articulated by Dr John Platt, that scientists should be the cartographers of the future, laying out for their fellow citizens a map of what was possible so that human societies could decide on which destinations and routes they wished to take Just over ten years ago, molecular biologists meeting at Asilomar seemingly echoed Dr Johnson's philosopher in accepting a voluntary moratorium on research in genetic engineering, a breathing space in which to consider the social implications To go the whole way with Dr Johnson would be wrong, however Quite apart from the impractibility of exercising the responsibility to suppress the findings of scientific work in a consistent way, the scientific estate, however much more knowledgeable than the rest of the population, has no right to take such decisions on behalf of society

The plain facts are that the decisions about science policy must in the end be taken as part of the normal political process and that much of the scientific enterprise of any country must be managed by the government of the day, if only because it is the main paymaster In the first chapter of his book, Dr Maurice Goldsmith tells the reader that societies will not be equipped to adopt a sensible attitude to contemporary science and scientists, let alone optimize science policy, unless there exists a body of "science critics" The science critic is defined as "a public policy journalist, alerting us to the growing pains of future worlds through the day-to-day discoveries of the present" No one would disagree with this proposition, but I doubt whether it is a sufficient as well as a necessary condition, societies are more complex than that and we know too little about their workings

Chapter 1 brings us to this conclusion via a great deal of verbiage and quotations from gurus such as C P Snow, C H Waddington, J D Bernal and Aldous Huxley Delightful they are too but quite unnecessary to the argument The remaining chapters, though less loaded with quotations, are refreshingly different They contain facts and tell us a good deal about scientific journalism, how it works, its priesthood and prizes They offer a good deal of practical advice to anyone thinking of taking up this profession and possibly, though I am in no position to judge, will be helpful to existing practitioners

The second and longest chapter is a description of how the various media operate in the Western world, and will repay attention by scientists who so often fail to get their message across and to understand the adaptations they must make to be successful in communicating to the public There follows a brief essay on science fiction, a subject in which I could never be interested, because it seemed to me that science itself was so much more interesting and called for much greater leaps of imagination Next there are two succinct essays on school science education (largely in Britain) and on how the message may be handled in a Third World country I greatly enjoyed both of them, the first because I knew something about the subject, the second because I did not and the confidence engendered by the sound and balanced presentation of the first spilled over and gave me confidence that the second must be of similar quality

Finally Dr Goldsmith returns to his new breed of animal, the science critic, whom he sees not just as a reporter of science but as someone who makes value judgements about it and its place in society But to do so such a person would have to be equipped with a broad synoptic view of science, be able to see the future and discern similarities, and be able to interpret and communicate science and uphold its integrity I could not but help feel that this chapter described the remarkable science teacher I had at school 60 years ago, which in turn made me consider the problem of generating a supply of such remarkable people to perform this necessary function This is briefly touched on by Dr Goldsmith, but I am sure he would be the first to admit that the time when we shall have an adequate supply of these critic-communicators is far distant The remaining 60 per cent of the book consists of two appendices, the first of which is concerned with some of the nuts and bolts of science journalism, and the second with examples of popular science writing

I enjoyed reading this book and I think many others would too The trouble is that it touches upon many topics, none of which is explored in any depth, and the main conclusions seem to be so miniscule Glancing at it again after my first reading, it seems to me that it was probably written, if not in Dr Johnson's record-breaking time, then very quickly from pieces that had already been published or were lying about nearly finished on the author's desk No criticism of Dr Goldsmith for that, but merely to say that whilst it is a good read it is not a great book ☐

Lord Dainton is Chancellor of the University of Sheffield, Sheffield S10 2TN, UK, and was, among other appointments, formerly President of the Association for Science Education and Chairman of the British Library Board

---

Going into a trance

Anthony W Clare

The Induction of Hypnosis. By William E Edmonston, Jr *Wiley 1986 Pp 432 £42 25, $43 95*

THERE are those, most notably Leon Chertok, who argue that the therapeutic element in psychoanalysis, in so far as there is one, is suggestibility, and that psychoanalysis, through its intermediary the transference, has not disengaged from hypnosis but has merely changed the terminology Freud himself declared in 1921 that "we still have no explanation of the nature of suggestion itself, that is to say of the conditions in which one is subject to its influence" It is indeed the case that ever since then many analysts have retained a lively interest in the subject of hypnosis

Contemporary hypnotists stress the importance of eye fixation and the attention-rivetting techniques employed to bring about the hypnotic state They do so with the same degree of purist vigour as analysts emphasize the crucial importance of free association and interpretation, and behaviourists insist on the essential therapeutic features of conditioning and deconditioning The real question, however, is the emotional aspects of the therapeutic relationship, the affective content of the procedures rather than the actual theoretical procedures themselves, yet it is this affective component of therapy, bound up as it is in notions such as suggestibility and susceptibility, which is poorly studied and indeed rarely considered

The fact remains that it was the nineteenth-century developments in hypnotherapy that facilitated the growth of modern psychotherapy, and it does not seem too far-fetched to believe that in the unravelling of hypnosis lies the answer to the thorny question of the efficacy of psychotherapeutic procedures in general Such issues are not the subject matter of Edmonston's book any more than they seem to be the concern of many psychotherapists But the contents of the book make it abundantly plain that such a state of affairs cannot last

If the author of this detailed text on the subject is to be believed, the history of hypnotic induction is the history of medicine It is true that the induction of sleep has been a mainstay of therapeutic interventions since before Hippocrates References to what sound like hypnotic trance states can be found in Vedic medicine, in the Papyrus Ebers of Egyptian medicine fifteen centuries before Christ and in the healing rituals of the Irish Druids before the arrival of St Patrick Not surprisingly, in the revised history of hypnosis Mesmer ceases to be the pivotal figure he seemed and is deemed to be not so much a pioneer as a catalyst, giving hypnosis a vital nudge forward at an opportune time

Edmonston devotes a great deal of space to a historical review of the role of hypnosis and to an enumeration of the many, varied and elaborate ways of inducing a hypnotic trance But at times the list of treatments falling under the rubric of hypnosis becomes rather elastic It is doubtful, for example, that the methods of Valentine Greatrakes, a contemporary of William Harvey, were truly hypnotic, given that he relied upon the laying on of hands and the smearing of his patients at St Bartholomew's Hospital with spittle!

The real issue in contemporary hypnosis is not its historical roots, nor its induction methods It is the precise relationship between classical hypnosis, with its emphasis on the induction of a trance or sleep-like state, and the newer task-orientated forms of behaviour therapy with their emphasis on relaxation, concentration, the relief of anxiety and the induction of tranquillity A related question is to do with the relationship between the analytic transference, so stressed within psychodynamic psychotherapy, and the notion of suggestibility which lies at the heart of hypnotic methods These, however, are matters that are scarcely touched upon in this book ☐

Anthony W Clare is Professor in the Department of Psychological Medicine, St Bartholomew's Hospital Medical College, West Smithfield, London EC1A 7BE, UK

Our lady of radium

H W Paul

Marie Curie: A Life. By Françoise Giroud Translated by Lydia Davis *Holmes & Meier, New York 1986 Pp 291 $34 50*

ROENTGEN'S discovery of X rays in 1895 led directly to the discovery of radioactivity by Henri Becquerel Unlike X rays, radioactivity did not excite scientists or the public, no lead-lined knickers were sold to protect prudent women from its scrutiny Then, in 1898, Marie and Pierre Curie announced their discovery of polonium and radium, and a radium craze soon infected the world The public had not been so interested in science since the publication in 1859 of Darwin's book on the origin of species

Maria ("Manyad") Sklodowska, a polyglot positivistic Pole from Warsaw, went to the University of Paris in 1891 to study science Polish nirvana freedom from the Russian yoke, cheap education, even for women, the opportunity to work for the improvement of mankind through science, and freedom to discuss the Polish condition without the threat of arrest Marie was a brilliant student and immediately attracted the attention of Sorbonnard mandarins such as Paul Appell, Edmond Bouty and Gabriel Lippmann

In 1895 she married into the tribe of French physicists, a rather inbred group The marriage changed Pierre Curie's career, for he was seduced into collaborating with Marie on the scientific obsession of the day, radioactivity Working in a shed at the municipal school of industrial physics and chemistry, they made their famous discoveries International fame led to success in Paris Pierre was given a chair at the Sorbonne, to prevent his going to Geneva, and after his death in 1906 the chair went to Marie The science faculty's radical precedent of opening its ranks to women was not followed by the Académie des Sciences, which narrowly preferred physicist Edouard Branly of the Institut Catholique — a minor triumph for Catholicism over republicanism Nearly a century before, women interested in science had been excluded from the newly-created faculty on the grounds that men would be more interested in them than in the professors' lectures That must rank as one of the world'ss worst academic administrative decisions In her long career Marie Curie won two Nobel prizes, one, with Becquerel and Pierre, in physics, and one for herself alone in chemistry Amongst the many virtues of the Third Republic (1871–1940) was that it made possible, if improbable, a career in science for women

In this book, Françoise Giroud has given us a very touching portrait of Marie

Curie. The history of science is sometimes handled cavalierly and the political passages are no better: Bismarck's Polish policy is reduced to a misleading *boutade*, and Prussia is generously given five million instead of two-and-a-half million Poles. Yet there is a great deal of reliable information on science artfully woven into the biography.

Madame Giroud often goes to the heart of matters with rare skill. The cultural and especially the social contexts are also deftly sketched, with emphasis on the economic exploitation of women. Glimpses into the private lives of the austere Curies and their more sybaritic colleagues provide some relief from Marie's grim devotion to science. But Marie Curie is not famous because she knew Paderewski or visited Rodin's studio, or even for her love affair with Paul Langevin which is treated in exquisite detail by Madame Giroud. Langevin's wife made the scandal notoriously public, and Paul Appell, the desperate dean of the faculty of sciences, wanted to solve matters by encouraging Marie to return to Warsaw. The American physicist B.B. Boltwood, a collaborator of Rutherford's, was satisfied that he had been correct in classifying her as a detestable idiot, but generous American support for Marie's work indicates that his opinion was not widely held in the United States.

The work of the Curies gave French science an international attention it had not enjoyed since the national orchestration of Pasteur's genius. The glory was equally deserved, and dozens of foreign researchers came to work on radioactivity at the Radium Institute in the Latin Quarter's vast city of science. In 1925 Marie picked Fred Joliot as her laboratory assistant, and his subsequent marriage to her scientist daughter Irene produced another great team in French physics. Joliot went on to become the main figure in the development of nuclear energy in France. Nearly up to her death in 1934 Marie was a tireless international ambassador for French science and an effective propagandist at home, where Jean Perrin and his friends were preparing the ground for the post Second World War structures and functions of French science.

Mary Evans

In collaboration — Marie and Pierre Curie.

Madame Giroud's book is heavily dependent on works such as Eve Curie's study of her mother. But its dramatic style and theme of *Une Femme Honorable* (the French title) make it a satisfying night's reading. ☐

H.W. Paul is in the Department of History, University of Florida, Gainesville, Florida 32611, USA, and is author of From Knowledge to Power: The Rise of the Science Empire in France, 1860–1939 *(Cambridge University Press, 1985).*

Optimal ideas

Brian Charlesworth

Physiological Ecology of Animals: An Evolutionary Approach. By R.M. Sibly and P. Calow. *Blackwell Scientific: 1986. Pp.179. Hbk £26, $49; pbk £11.95, $24.50.*

OVER the past decade, the quantitative application of the concept that the properties of living organisms represent optimal solutions to specific design problems has been a growth industry. Richard Sibly and Peter Calow have been prominent exponents of this approach, and the aim of this book is to set out some principles of optimization theory, and to apply them to a series of problems in physiological ecology and life-history evolution. The basic procedure is to assume a plausible set of constraints acting upon the properties of interest in the particular study, for example that an increase in the amount of effort devoted to reproduction at a given stage of the life-cycle leads to a decrease in the chance of survival to the next stage. Given a criterion by which to measure darwinian fitness, one then calculates the set of values of the properties that yield a maximum in fitness according to the specified constraints. It is then assumed that natural selection will have caused populations of organisms to evolve close to this optimal set. Their observed properties are compared with the predictions of the model, in an effort to test the model's validity.

This procedure is often criticized on the grounds that failure of an optimization model to fit the real world does not provoke its originator to abandon belief in natural selection as the main force guiding the evolution of the characteristics of interest; rather, it usually leads him to tinker with the details of the model. This criticism is somewhat misplaced, because the aim of workers using these methods is not to test the general validity of the theory of natural selection, but to uncover the biological factors which may be of significance in moulding the action of selection on the traits in question. Furthermore, given the complexity of biological systems one would have to be wildly optimistic to expect an accurate fit between predicted and observed properties. So emphasis has often been placed on theoretical predictions of the consequences of variation in ecologically determined factors which modulate the ways in which the system responds to selection, and on comparisons of populations with appropriate differences in ecology to see if they differ (at least qualitatively) in the expected fashion. Thus, it is a stock-in-trade of life-history theory that species living in environments with high adult survival relative to juvenile survival should invest less in reproductive effort per breeding season, compared with otherwise similar species with low adult survival.

Sibly and Calow bring out these principles in a number of case studies, ranging from an analysis of the factors affecting the efficiency of guts, to strategies of growth and reproduction. The emphasis is on the development of optimization models, rather than detailed testing against the data; indeed they state that their aim is to stimulate experimentalists to provide the necessary data. They succeed in convincing one of the diversity of topics to which their approach can be applied, and the book will be of considerable use to biologists in this area.

The authors can, however, be criticized on several points. One is that they give little indication of previous contributions to life-history theory, a relatively mature area of enquiry. Instead, they present their own reformulations of classic models as though they were wholly original, and provide only scanty citation of the literature. This will not help beginners find their way. More importantly, they fail to mention alternatives to optimization methods. Thus, there is no reference to Williams's and Medawar's theories of senescence, which have an important bearing on life-history evolution. Little mention is made of the possible role of competition in moulding selection on growth and size, and Maynard Smith's Evolutionarily Stable Strategy method for handling the complications that arise in competitive situations is not discussed. While the authors' approach has considerable value, and they have every right to advocate it, their book cannot be said to offer a comprehensive survey of the methods needed to solve problems in evolutionary ecology. ☐

Brian Charlesworth is a Professor in the Department of Biology, University of Chicago, 1103 East 57th Street, Chicago, Illinois 60637, USA.

Steps out of time

Eric Ashby

Walkers. By Miles Jebb. *Constable, London: 1986. Pp.202. £10.95.*

SIXTY years ago, devotees of the non-fiction book at bedtime were well supplied by a genre which is now rare: the essay. Contemporaries of Chesterton, Belloc and Aldous Huxley had plenty of entertainment awaiting them at the end of the day. Today there are very few competent practitioners in that style of writing. One can still enjoy the classical essayists — Charles Lamb for the light-hearted, Emerson for the sober-minded, and Montaigne and Bacon — but who is writing in that style now?

Well, Miles Jebb for one. His book, *Walkers*, is in the tradition of the whimsical nineteenth-century essayists. In eleven brief chapters Jebb writes about the various kinds of people who walk or have walked: pilgrims, athletes, intellectuals, tramps, ramblers... I'll not disclose the whole list. The book is a compilation of anecdotes and episodes from history and fiction, told with the comfortable informality of after-dinner conversation, by a man with a well-stocked mind.

If you start to delve into the literature about walking, it is surprising what turns up. Each of Jebb's eleven kinds of walker finds himself heir to certain conventions. The value of a pilgrimage may be enhanced by discomfort: "The Countess of Clare threw away her shoes as she started on a pilgrimage". Athlete–walkers are out to break records: Sean Maguire, who "in 1978 walked the 7327 miles from Yukon to Florida in three hundred and seven days"; and Sebastian Snow, who reckoned he had taken some 11,745,000 paces in his walk from the tip of South America to Panama City with a 60 lb pack on his back. Intellectual walkers are the most articulate (witness Hilaire Belloc's *Path to Rome*) and they often philosophize as they walk. Or they boast of a kind of muscular intellectualism: thus Leslie Stephen, when at Cambridge, walked to London (over 50 miles) in 12 hours to attend a dinner of the Alpine Club.

Tramps fall into two categories: those who walk because they have nowhere else to go, and those who "walk away from convention". The former are inarticulate (though the poet W.H. Davies was an exception); the latter, such as Stevenson and Maxim Gorky, often become tramps in order to collect material for writing. Ramblers not only ramble; they combine to form pressure groups to lobby for conservation, and they sometimes flock together in group walking. They are somewhat querulous, having been driven off the roads by traffic, but we are indebted to them for their lobbying: in the United States, the Sierra Club has a fine record in defending the wilderness, and in Britain the Ramblers' Association has preserved many footpaths that would otherwise have been obliterated by farmers.

If you like erudite and genial conversation about themes such as these, this is a bedtime book to savour. ☐

Lord Ashby is a Fellow of Clare College, Cambridge CB2 1TL, UK.

This is a great book,
and an exciting book...
readable, worth reading
and enlightening.
Sir Karl Popper

Cosmology,
Physics &
Philosophy
by
Prof Ben Gal-Or

"a master piece... any good library must have a copy of this classical work."
Ind. J. Phys.

"... a tour de force... will be widely read and appreciated... a magnificent and sustained piece of work!"
Sir Alen Cottrell

... remarkable... challenging...
Am. J. Phys.

2nd Printing
Springer-Verlag
175 5th Ave., N.Y., 10010

Reader Service No.7

Thought of biology

P. T. Saunders

Biophilosophy: Analytic and Holistic Perspectives. By Rolf Sattler. *Springer-Verlag: 1986. Pp.284. Pbk DM66*

ROLF Sattler is a biologist who believes that it is important to understand the fundamental presuppositions of biology and how they influence research and even our view of life itself. In *Biophilosophy* he sets out to introduce other biologists to these ideas.

I am in full sympathy with Sattler's aim, and he has gone some way towards achieving it. But there are reservations. The book seems to be intended as both an introductory text and a monograph, and rather falls between the two stools. Far too much material has been crammed into a text that also has to include elementary expositions of basic ideas. As a result, some explanations are marred by what are, for the beginner, digressions. At other times, Sattler rushes along at great speed; he deals with the issue of free will in just over a page. Many terms and ideas are mentioned so fleetingly that the reader not only is left uninformed about them but will not even be able to grasp how they are relevant to the discussion.

Often Sattler does no more than mention that a certain idea has been put forward and give some references. For example, he considers that the idea of species as individuals rather than classes leads to the scientific view of unity or oneness to which he devotes the last chapter. Yet he says very little in support of the idea, beyond mentioning the common gene pool and providing some references to "a critical appraisal" of the concept. It would have been better to spend more time on those questions that he considers to be most important, and to compensate by omitting some other material altogether.

Biophilosophy, then, is not a satisfying book. The author seems to have tried to get down in print everything he has thought about, from plant morphology to Zen Buddhism. Further, aspects of the selection of material and the relative emphasis given to different topics mean that he has not fully assimilated all that he has learned into a coherent whole. On the other hand, if *Biophilosophy* loses by not having been written by a professional philosopher, it also gains from having been written by a working scientist to whom these ideas matter, and who sees many of the issues from the same point of view as his readers.

Far too few biologists are aware of the philosophical issues that underlie and influence what they do. Moreover, because so much of the philosophy of science is concerned with physics, not biology, it is not readily accessible to them. Biologists who read this book will become acquainted with many important questions, and, what is more, will have them explained in familiar terms — there are not many introductions to the philosophy of science that illustrate the basic concepts through plant morphology, instead of quantum mechanics. I even approve of reminding the reader that plant morphology and Zen Buddhism are both worth thinking about, and not necessarily in separate intellectual compartments. I only wish Sattler had tightened the whole thing up a bit. ☐

P.T. Saunders is in the Department of Mathematics, King's College London (KQC), Strand, London WC2R 2LS, UK.

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse. I am interested in electronic ordering through the following system(s):

☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem
☐ Other (please specify)________
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name________
Title________
Institution/Company________
Department________
Address________
City________State____Zip________
Phone (____)________

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

SCIENCE AND TECHNOLOGY

Whatever the subject... refer to MACMILLAN DICTIONARIES

DESIGNED FOR STUDENTS

The paperback *Macmillan Dictionary Series* combines new and established works of teaching and reference with an attractive series format.

- ∗ authoritative definitions and explanations of all current terminology written by experts
- ∗ lucidly written entries with thorough cross referencing
- ∗ regular new editions keep the student up to date with the latest research, practice and technology
- ∗ a clear layout for easy reference
- ∗ prices range from just £7.95 to £9.95

NEW IN 1986...

MACMILLAN DICTIONARY OF BIOTECHNOLOGY
JIM COOMBS March 1986 320pp 0 333 40875 6 £9.95

MACMILLAN DICTIONARY OF THE HISTORY OF SCIENCE
Edited by W. F. BYNUM, E. J. BROWNE and ROY PORTER
March 1986 528pp 0 333 34901 6 £8.95

MACMILLAN DICTIONARY OF PERSONAL COMPUTING AND COMMUNICATIONS DENNIS LONGLEY and MICHAEL SHAIN
August 1986 320pp 0 333 42170 1 £8.95

MACMILLAN DICTIONARY OF PHYSICS
MARY P. LORD August 1986 320pp 0 333 42377 1 £8.95

MACMILLAN DICTIONARY OF IMMUNOLOGY
FRED ROSEN, LISA STEINER and EMIL UNANUE
August 1986 256pp 0 333 39248 5 £8.95

RECENTLY PUBLISHED ...

MACMILLAN DICTIONARY OF LIFE SCIENCES *Second Edition*
E. A. MARTIN 1985 400pp 0 333 38647 7 £9.95

MACMILLAN DICTIONARY OF THE ENVIRONMENT
Second Edition
MICHAEL ALLABY 1985 536pp 0 333 37951 9 £8.95

MACMILLAN DICTIONARY OF INFORMATION TECHNOLOGY
Second Edition DENNIS LONGLEY and MICHAEL SHAIN
1985 392pp 0 333 37261 1 £8.95

MACMILLAN DICTIONARY OF MICROCOMPUTING *Third Edition*
CHARLES SIPPL 1985 544pp 0 333 37082 1 £8.95

MACMILLAN DICTIONARY OF DATA COMMUNICATIONS
Second Edition
CHARLES SIPPL 1985 544pp 0 333 37083 X £8.95

MACMILLAN DICTIONARY OF ENERGY
Edited by MALCOLM SLESSER
1985 314pp 0 333 36867 3 £7.95

MACMILLAN DICTIONARY OF ASTRONOMY
VALERIE ILLINGWORTH 1985 320pp 0 333 39243 4 £9.95

Not available in USA.

M MACMILLAN PRESS

Macmillan Press, Houndmills, Basingstoke, Hants RG21 2XS, UK

CONVENIENT·INFORMATIVE·TIMELY·PROVOCATIVE

FOUR DYNAMITE REASONS WHY YOU SHOULD SUBSCRIBE TO BIO/TECHNOLOGY NOW . . .

1. You'll save time. Because you don't have time to read every journal in the field, BIO/TECHNOLOGY is written and edited to give you all the information you need in ONE convenient source. BIO/TECHNOLOGY is the only journal to give you news. Features. Original research. Product evaluations. Industrial applications. Job openings. Book and product reviews. And more. If it's important to your work, you'll find it between the covers of BIO/TECHNOLOGY.

2. You'll learn new ways to improve your on-the-job skills and run a more successful operation. Each month BIO/TECHNOLOGY gives you the scoop on invaluable start-up information . . . the latest instrumentation . . . marketing breakthroughs . . . funding sources . . . new technologies . . . patents and licences . . . profiles on how others are avoiding the pitfalls and reaping the rewards. In fact, just one practical tip from BIO/TECHNOLOGY can be worth many times the cost of your subscription.

3. You'll be up-to-date on all major industry developments. With correspondents all over the world, BIO/TECHNOLOGY delivers the inside story on who's doing what, where. You'll get the latest word on changing government regulations and tax law, and how they affect your work.

Plus you'll read original research papers by bench scientists, engineers, and biologists that will help keep you in the forefront of every new discovery worth knowing about.

4. You'll get new ideas you can put to use immediately. BIO/TECHNOLOGY lets you rub elbows with eminent scientists, tour new facilities and learn how new processes work in every field from genetic engineering to molecular biology. Each month, BIO/TECHNOLOGY helps spark ideas for your own research, for new industrial applications and for getting along in the increasingly competitive financial arena.

If you're not already subscribing to BIO/TECHNOLOGY, don't delay the moment any longer - return the coupon below and we will send you twelve issues. You risk nothing. If for any reason BIO/TECHNOLOGY doesn't live up to your highest expectations, just let us know and you'll receive a full refund on all unmailed copies. Remember too, that your subscription may be tax deductible.

TO START RECEIVING YOUR SUBSCRIPTION, FILL OUT THE COUPON BELOW AND RETURN WITH PAYMENT TO:

U.S.A.: BIO/TECHNOLOGY
Subscription Order Dept., P.O. Box 1543
Neptune, New Jersey 07754-1543 U.S.A.

To speed up your subscription in the U.S.A., call toll free 1-800-524-0328.

ENGLAND: BIO/TECHNOLOGY
Subscription Order Dept., Macmillan Journals, Ltd.
Brunel Road, Houndmills, Basingstoke
Hampshire RG21 2XS, England

☐ YES! Please send me one year of BIO/TECHNOLOGY (12 issues). I understand that I may cancel my subscription at any time and receive a full refund on all unmailed issues.

☐ U.S.A.-$59.00* ☐ Australia-A$130**
☐ U.K.-£70** ☐ Elsewhere-U.S.$112**

Japan Publications Trading Company Ltd
2-1 Sarugaku-cho 1-chome
Chiyoda-ku, Tokyo, Japan Tel: (03) 292 3755

Note: *Represents 24% off Basic Subscription Price of U.S. $78.00
**All other prices quoted are "Air Speed"

☐ My payment is enclosed at the price checked.
☐ Please charge my credit card account:
☐ American Express ☐ Mastercard ☐ Visa

Card No.: ______________________ Exp. Date: __________
Signature: ______________________
Name: ______________________
Company: ______________________
Address: ______________________
City/State/Postal Code: ______________________
Country: ______________________

7AUS

Forty years of genetic recombination in bacteria

Between April and June 1946, Joshua Lederberg and Edward L. Tatum carried out a series of experiments that proved that bacteria can exchange their genes by sexual crossings. The experiments were reported in Nature *just 40 years ago[1]. In the following pair of articles, Joshua Lederberg first provides a personal reminiscence of the circumstances of the discovery and then, together with Harriet Zuckerman, considers it as a possible case of 'postmature' scientific discovery.*

A fortieth anniversary reminiscence

Joshua Lederberg

In September 1941, when I started as an undergraduate at Columbia University, the genetics of bacteria was still a no-man's-land between the disciplines of genetics and (medical) bacteriology. The question whether "bacteria have genes, like all other organisms" was still unanswered, indeed rarely asked. My own thoughts at that moment lay elsewhere. I looked forward to a career in medical research applying chemical analysis to problems like cancer and the malfunctions of the brain. Cytotoxicology then appeared to be the most promising approach to cell biochemistry. It was Frances J. Ryan (d. 1963) who turned my attention to the sharper tools of genetics.

Ryan had spent 1941–42 as a postdoctoral fellow at Stanford University, where he had met G. W. Beadle and E. L. Tatum (d. 1975), and had become fascinated with their recent invention of nutritional mutations in *Neurospora* as a tool for biochemical genetic analysis[2]. Although working on a fungus like *Neurospora* did not go down smoothly in a Department of Zoology as at Columbia, where Ryan had accepted an instructorship, he established a laboratory to continue these studies. In January 1943 I was fortunate to get a job in his laboratory assisting in the preparation of media and handling of *Neurospora* cultures. Ryan's personal qualities as a teacher and the setting of serious research, discussion with him, other faculty members and graduate students in the department nourished my education as a scientist. On 1 July 1943, I was called to active duty in the United States Naval Reserve, and my further months at Columbia College alternated with spells of duty at the United States Naval Hospital, St Albans, Long Island. There, in the clinical parasitology laboratory, I had abundant opportunity to observe the life cycle of *Plasmodium viva*. This experience dramatized the sexual stages of the malaria parasite, which undoubtedly sensitized me to the possibility of cryptic sexual stages in other microbes (perhaps even bacteria). In October 1944, I was reassigned to begin my studies at Columbia Medical School; but I continued working with Ryan at the Morningside Heights campus.

Lederberg in 1945

Discovery

The important biological discovery of that year, by Avery, MacLeod and McCarty, was the identification of DNA as the substance responsible for the *Pneumococcus* tranformation[3]. This phenomenon could be viewed as the transmission of a gene from one bacterial cell to another; but such an interpretation was inevitably clouded by the obscure understanding of bacterial genetics at the time. Avery's work, at the Rockefeller Institute in New York, was promptly communicated to Columbia biologists by Theodosius Dobzhansky (who visited Rockefeller) and by Alfred Mirsky (of the Rockefeller faculty) who was a close collaborator of Arthur Pollister in the Zoology Department. The work was the focus of widespread and critical discussion among the faculty and students. Mirsky was a vocal critic of the purported chemical identification of the transforming agent, while applauding the central importance of the work. For my own part, the transcendent leap was simply the feasibility of knowing the chemistry of the gene. Whether this was DNA or protein would certainly be clarified quickly, provided the *Pneumococcus* transformation could be securely retained within the conceptual domain of gene transmission. I read the Avery, MacLeod and McCarty paper on 20 January 1945, prompted by Harriett Taylor (later Ephrussi-Taylor) a graduate student in Zoology who planned to pursue her postdoctoral studies with Avery. My excited response is recorded as ... "unlimited in its implications... Direct demonstration of the multiplication of transforming factor... Viruses are gene-type compounds."

At once, I thought of attempting similar transformations by DNA *in Neurospora*. This organism had a well understood lifecycle and genetic structure. The biochemical mutants opened up by Beadle and Tatum also allowed the efficient detection of nutrionally self-sufficient (prototrophic) forms, even if these were vanishingly rare. This would facilitate the assay of transformational events.

Between January and May, 1945, I shared this idea with Francis Ryan; in June, he invited me to work on the subject with him. To our dismay, we soon discovered that the leucine-minus *Neurospora* mutant would spontaneously revert to prototrophy[4], leaving us with no reliable assay for the effect of DNA in mediating genetic change in *Neurospora*. Questions about the biology of transformation would remain inaccessible to conventional genetic analysis if bacteria lacked a sexual stage. But was it true that bacteria were asexual? Rene Dubos' monograph, *The Bacterial Cell*[5], footnoted how inconclusive the claims were for or against any morphological exhibition of sexual union between bacterial cells.

My notes dated 8 July 1945 detail hypothetical experiments both to search for mating among *Monilia* (medically important yeast-like fungi) and to seek genetic recombination in bacteria (by the protocol that later proved to be successful). These notes coincide with the beginning of my course in medical bacteriology. They were provoked by the contrast of the traditional teaching that bacteria were *Schizomycetes*, asexual primitive plants, with an appreciation of sexuality in yeast[6], which was represented at Columbia by the graduate research work of Sol Spiegelman and Harriett Taylor.

Dubos[5] cited many unclear, and two clear-cut negative results[7,8] for sexuality in bacteria using genetic exchange methodology. But these two studies had no selective method for the detection of recombination and so would have overlooked the process had it occurred less often than perhaps once per thousand cells. With the use of a pair of nutritional mutants, say A^+B^- and A^-B^+, one could plate out innumerable cells in a selective medium and find a single A^+B^+ recombinant. In early July, I began experiments along these lines. In the first instance I used a set of biochemical mutants in *Escherichia coli*, which I began to accumulate in Ryan's laboratory. To avoid the difficulty that had arisen in our *Neurospora* experiments, a spontaneous reversion from A^-B^+ to A^+B^+, the strategy would be to use a pair of double mutants: $A^-B^-C^+D^+$ and $A^+B^+C^-D^-$. Sexual crossing should still generate $A^+B^+C^+D^+$ prototroph recombinants. These would be unlikely to arise by spontaneous reversions which, in theory, requires the coincidence of two rare events; $A^- \rightarrow A^+$ and $B^- \rightarrow B^+$. Much effort was devoted to control experiments to show that double reversions would follow this model, and so occur at a negligible frequency in the cultures handled separately. Thus the occurrence of prototrophs in the mixed cultures would be presumptive evidence of genetic recombination.

Long shot

Meanwhile at Stanford, Ed Tatum, whose doctoral training at Wisconsin had been in the biochemistry of bacteria, was returning to bacteria as experimental subjects, having published two papers on the production of biochemical mutants in *E. coli*[9], including double mutants like those described here. During the summer of 1945 Francis Ryan learned that Tatum was leaving Stanford to set up a new programme in microbiology at Yale. He suggested that, rather than merely ask Tatum to share these new strains, I apply to work with him and get the further benefit of his detailed experience and general wisdom. Tatum agreed and suggested that I arrive in New Haven in late March, to give him time to set up his laboratory. He hinted that he had some similar ideas of his own, but never elaborated them. The arrangement suited him by leaving him free to complete his work on the biochemistry of *Neurospora*, perform the heavy administrative duties of his new programme, and still participate in the long-shot gamble of looking for bacterial sex.

Experimental luck

1. We have learned[12] that *E. coli* strain K-12 itself was a remarkably lucky choice of experimental material: only about one in twenty randomly chosen strains of *E. coli* would have given positive results in experiments designed according to our protocols. In particular, strain B, which has become the standard material for work on bacteriophage, would have been stubbornly unfruitful. Tatum had acquired K-12 from the routine stock culture collection in Stanford's microbiology department when he sought an *E. coli* strain to use as a source of tryptophanase in work on tryptophan synthesis in *Neurospora*[13]. The same strain was then in hand when he set out to make single, and then double mutants in *E. coli*[9]. In 1946, I was very much aware of strain specificities and was speculating about mating types (as in *Neurospora*). I have no way to say how many other strains would have been tried, or in how many combinations, had the June 1946 experiments not been successful.

2. An equally important piece of luck was that, the selected markers Thr (threonine) and Leu (leucine) are found almost at the origin of the *E. coli* chromosome map[14]. The cognoscenti will recognize that in a cross $B^-M^-T^+L^+F^+ \times B^+M^+T^-L^-F^-$, the configuration used in June 1946, these chromosome localizations offer almost a maximum yield of selectible recombinants. We were therefore led stepwise into the complexities of mapping.

It took about six weeks, from the first serious efforts at crossing in mid-April 1946, to establish well-controlled, positive results. These experiments could be done overnight, so the month of June allowed over a dozen repetitions, and the recruitment of almost a dozen genetic markers in different crosses. Besides the appearance of $A^+B^+C^+D^+$ prototrophs, it was important to show that additional unselected markers in the parent stocks would segregate and recombine freely in the prototrophic progeny. This result left little doubt as to the interpretation of the experiments.

An immediate opportunity for public announcement presented itself at the international Cold Spring Harbor Symposium in July. This was dedicated to the genetics of microorganisms, signalling the postwar resumption of major research in a field that had been invigorated by the new discoveries with *Neurospora*, phage, and the role of DNA in the *Pneumococcus* transformation. Tatum was already scheduled to talk about his work on *Neurospora*. We were granted a last-minute improvisation in the schedule to permit a brief discussion of our new results.

The discussion was lively. The most principled criticism came from Andre Lwoff who worried about cross-feeding of nutrients between the two strains without their having in fact exchanged genetic information. Having taken great pains to control this possibility, I felt that the indirect genetic evidence was quite conclusive. Fortunately, Max Zelle mediated the debate, and generously offered to advise and assist me in the direct isolation of single cells under the microscope. These subsequent observations did quiet remaining concerns of the group that Lwoff had assembled at the Pasteur Institute, including Jacques Monod, Francois Jacob and Elie Wollman, who were to make the most extraordinary contributions to the further development of the field. The single cell methods were also useful in later investigations in several fields. A direct result of the Cold Spring Harbor meeting was the prompt ventilation of all the controversial issues. With a few understandable, but minor points of resistance, genetic recombination in bacteria was soon incorporated into the mainstream of the burgeoning research in molecular biology, and after another decade or so into the standard texts of bacteriology. It still took some years to work out the intimate details of crossing in *E. coli*; some, including the crucial question of the physical mechanism of DNA transfer between mating cells, are still obscure.

The public image of the scientific fraternity today has seldom been so problematic and the system cannot avoid putting a high premium on competition and self-assertion. We can recall with gratification how the personalities of Ryan[10] and Tatum[11] exemplified norms of nurture, dignity, respect for others, and above all a regard for the advance of knowledge.

Joshua Lederberg is at the Rockefeller University, New York, New York 10021. The research summarized in this article was supported in 1946 by a fellowship of the Jane Coffin Childs Fund for Medical Research.

1. Lederberg, J. & Tatum, E.L. *Nature* **158**, 558 (1946).
2. Beadle, G.W. & Tatum, E.L. *Proc. natn. Acad. Sci. U.S.A.* **27**, 499 - 506 (1941).
3. Avery, O.T., MacLeod, C.M. & McCarty, M. *J. exp. Med.* **79**, 137 - 158 (1944).
4. Ryan, F.J. & Lederberg, J. *Proc. natn. Acad. Sci. U.S.A.* **32**, 163-173 (1946).
5. Dubos, R. *The Bacterial Cell* (Harvard, Cambridge, 1945).
6. Winge & Lausten, O. *C.R. Lab. Carlsberg, Ser, physiol.*, **22**, 99-119, (1937).
7. Sherman, J.M. & Wing, H.U. *J. Bact.* **33**, 315 - 321 (1937).
8. Gowen, J.W. & Lincoln, R.E. *J. Bact.* **44**, 551 - 554 (1942).
9. Gray C.H. & Tatum, E.L. *Proc. natn. Acad. Sci. U.S.A.* **30**, 404-410 (1944).
10. Lederberg, J. *in University on the Heights*, (ed. First, W.) 105 - 109. (Doubleday, Garden City, New York, 1969).
11. Lederberg, J. *A. Rev. Genet.* **13**, 1-5 (1979).
12. Lederberg J. *Science* **114**, 68 - 69 (1951).
13. Tatum, E.L. & Bonner, D.M. *Proc. natn. Acad. Sci. U.S.A.* **30**, 30 - 37 (1944).
14. Bachmann, B.J. *Microb. Revs.* **47**, 180 - 230 (1983).

Postmature scientific discovery?

Harriet Zuckerman and Joshua Lederberg

NEW scientific discoveries do not always flow directly from those made just before Rather, several varieties of discontinuity can be identified in the growth of science Premature discoveries are those that scientists do not attend to in a timely way, and are retrospectively described as having been "ahead of their time" These have been examined by Barber[1] and Stent[2] Here, we suggest that there are also postmature discoveries, those which, are judged retrospectively to have been 'delayed' We analyse the arguments that the discovery of bacterial sex was postmature and take up the correlative questions of how the problem was identified, and why Lederberg and Tatum[3,4] were likely candidates for making it when they did

This paper draws on documents, published and private, and analyses by the sociologist-observer and the scientist-participant Our dialectic procedure departs from most oral histories[5,6], first, the procedure was iterative as new discussion raised further possibilities, we both searched for relevant documentation, and second, we both identified the underlying analytic questions and articulated tentative answers to them We felt that personal reminiscence had to be validated by contemporary documents and other testimony as oral history and autobiography are prone to "unconscious falsification"[7]

Continuities and discontinuities

Scientific growth, usually broadly incremental, can at important times be episodic and discontinuous Premature discoveries, one conspicuous form of temporal discontinuity in science, are either passively neglected or actively resisted at the time they are made Mendel's discovery of particulate inheritance in 1865, lost to view for thirty-five years, is the best-known historical case Discoveries can be premature because they are conceptually misconnected with 'canonical knowledge'[2], are made by an obscure discoverer, published in an obscure place, or are incompatible with prevailing religious and political doctrine Barriers between disciplines imposed by specialization of inquiry also contribute to neglect or resistance[1,8-10] Although the character and sources of premature discoveries have received some analytical attention[8,11], the pattern of postmature discoveries has not been identified, much less systematically studied

For a discovery to qualify as postmature, for it to evoke surprise from the pertinent scientific community that it was not made earlier, it must have three attributes In retrospect, it must be judged to have been technically achievable at an earlier time with methods then available It must be judged to have been understandable, capable of being expressed in terms comprehensive to working scientists at the time, and its implications must have been capable of having been appreciated

Both prematurity and postmaturity can be recognized only by retrospection They differ in that prematurity is a matter of actual historical observation while postmaturity is a matter of retrospective conjecture Such formulations would seem to smack of 'Whig History', the inclination, according to Butterfield, "to produce a story which is the ratification if not the glorification of the present"[12] But, they are designed to serve quite the contrary purpose The ideas of premature and postmature discovery provide convenient handles for analysing discontinuities in the growth of scientific knowledge, and support a nonlinear and complex model of advancement in scientific understanding

Postmature discoveries are not all of a piece One class results from pre-emption of scientists' research attention For example, Linus Pauling observed that there was "no reason why" he, himself, could not have discovered the alpha helix eleven years earlier than he actually did "after a few hours of work" In fact, he was preoccupied in the interval by other seemingly more important and feasible inquiries[13,14] Another class of postmature discoveries answer questions not previously recognized by scientists to be problematic Certain assumptions, beliefs and images[15] which are also indispensable for the organization of scientific thought can, in specific cases, impede perception of lines of inquiry For example Weinberg notes that physicists neglected to pursue quantum field theory further in the 1930s because prevailing images, conceptual schemes and attitudes toward theory and empirical evidence stood in the way[16] In our case study, both cognitive and social processes obstructed the thinking of scientists about recombination in bacteria

Sources of neglect

Why was recombination in bacteria not perceived as problematic before 1946? How had asexuality in bacteria come to be an unquestioned 'truth' and how was that view perpetuated?

Before 1870, many believed that the different shapes bacteria assumed were varieties of the same organism, which changed under varying conditions Indeed, the doctrine of polymorphism or bacterial plasticity became the basis for extravagant claims about variability through most of the nineteenth century By 1872, Ferdinand Cohn concluded that the various shapes bacteria took were not different forms of the same organism, they were monomorphic and did not change during their short lifetimes[17] Yet reports of variation continued until 1881 when Robert Koch introduced a simple and effective means for growing pure cultures Koch's pure-culture method, which became a symbol of modern bacteriology with its phobia of contamination, together with Cohn's doctrine of monomorphism rapidly changed bacteriologists' views about variation The two were consolidated into what was called the Cohn–Koch Dogma, which discouraged for years the study of the problems of morphology, inheritance and variation in bacteria[18]

Cohn was convinced that bacteria were primitive plants which could "only reproduce by asexual means" and in 1875 characterized all bacteria as *Schizomycetes* or 'fission fungi' With every use of that label, bacteriologists were reminded that these organisms reproduced only by fission and that they were simple primitive plants, a tradition that had begun with Leeuwenhoek's first observation of bacteria in 1675 Labels, categories, nomenclature and taxonomies usually help to organize scientific thought but can also delay the reexamination of fallacious traditions, thus becoming self-fulfilling prophecies[19] In the end, the emergence of medical microbiology as a science depended on the doctrinal base laid down by Cohn and the pure culture methods of Koch Nonetheless, monomorphic doctrine, when strictly construed, threw out the baby of bacterial variation with the dirty bath water of contamination It was widely assumed that observations of bacterial variation had to result from contamination Bacteriologists took experiments involving variation to be error-prone and disreputable[20] Such experiments were to be avoided as having great procedural difficulty and little intellectual merit With the strong incentives in science for avoiding problems notorious for leading to irreproducible results, very few scientists would elect to undertake them

Bacteria occupied an ambiguous place in the hierarchy of living organisms To many, these organisms appeared so primitive that they could not yet have evolved 'differentiated genes' This image also reinforced the use of bacteria as exemplars of pre-genic levels of organization for physico-chemical analysis Once such complex imagery becomes established, special provocation is needed to splinter away one or more of its elements

Disciplinary emphases and the division

of labour among the sciences also diverted attention from the problem of bacterial sexuality. Bacteriologists were principally concerned with problems in medical pathology rather than issues like the biology of bacterial reproduction. Geneticists were no more interested in bacterial reproduction than bacteriologists. They were occupied with larger organisms in which the products of crossing were readily observed. Thus, disciplinary division of labour and the careful choice of organisms for inquiry, both generally conducive to the development of scientific knowledge, contributed in this instance to neglect of bacterial recombination. It has been argued, however, that 'disciplinary dogmatism' and 'disciplinary monopoly' have only rarely impeded the development and diffusion of scientific innovation[21].

Members of the Delft School of Microbiology, in the early part of this century, did bridge the gap between bacteriology and genetics. Clearly separating themselves from the medical bacteriologists who maligned bacteria, they believed that progress in fundamental microbiology depended on people who 'loved' microbes[22]. Martinus Beijerinck, the main figure in the group, seems now to have been the most likely candidate for investigating bacterial sex. He rejected prevailing dogma on bacterial invariability, promptly sited deVries' finding on plant mutations and offered some of the first coherent challenges to strict monomorphism[23]. He also developed 'enrichment culture' methods, forerunners of the selective techniques used later in discovering bacterial recombination. Moreover, he was better informed than most microbiologists about work on plant hybridization which would have been useful in planning any investigation of sex in bacteria. Beijerinck and the Delft School were likely candidates for investigating bacterial sex, but they did not. In fact, Beijerinck strongly supported the Cohnian dogma of Schizomycetes. Thus the problem of sexual recombination still fell between disciplinary schools[24].

Significance of bacterial sex

By the 1930s, developments were under way that led biologists to reexamine how bacteria related to other forms of life and whether bacteria really had genes. Important among these developments was the unification in biological thought of Mendelian genetics, quantitative population theory and darwinian evolution, particularly the notion of species being Mendelian breeding populations or isolated gene pools. The idea that sexuality was, itself, an evolved genetic system proved particularly provocative, with illustrations drawn from simple and complex plant life. Dobzhansky's monograph, "*Genetics and the Origin of Species*"[25], was widely read as the definitive reinterpretation of darwinian theory of evolution and focused interest on the details of breeding systems as the key to understanding evolutionary development. This, in turn, sharpened interest in understanding the evolution of organisms, like bacteria, believed to be devoid of sexual mechanisms.

The biochemical analysis of microbial nutrition, especially by Knight and Lwoff[18], was another major impetus to reexamining the relationship of bacteria to other forms of life. In particular, the discoveries that the biochemistry of microbes paralleled in many details that of higher organisms inspired Beadle and Tatum's work on *Neurospora* in 1941[26]. They showed *Neurospora's* usefulness as a research material for studying the genetic control of an organism's development through the encoding of specific enzymes, known as the 'one-gene-one-enzyme' hypothesis. This marriage of biochemistry and genetics and particular significance for the Lederberg–Tatum work[27,28].

There was also renewed speculative interest in a biochemical theory of the origin of life. "*The Origin of Life on Earth*"[29] by the Russian biochemist, Oparin, became available in English in 1944, as did "*What is Life?*"[30] by the physicst, Schrodinger. Both focused attention on questions that demanded the integration of the biology of viruses and microbes with the more traditional biology of plants and animals.

The connections between these independent developments were not always apparent at the time. But one event did call attention to their common message, the discovery by Avery, MacLeod and McCarty in 1944 which identified DNA as the transforming principle that changed rough non-pathogenic pneumococci into smooth virulent ones.

The scientific significance of that discovery has been examined in detail[32-38]. For our purposes, it highlighted two important questions: what was the structure of bacterial genes and how were they transferred? Thus the work by Avery *et al* made the question of bacterial sex newly consequential. Dubos[39] makes it clear that had sexual reproduction been observed, it would have been understood and appreciated. But bacteria were so widely assumed not to reproduce sexually that noone considered this problem to be important. Dogma prevailed over focused curiosity.

Structural contexts

In retrospect, Lederberg's position in the communication network and his not yet having a career in science seem consequential for his identifying the problem of bacterial sex, for his developing a method for its investigation and for his being in a position to do the research. Tatum was led to the problem independently for somewhat different reasons[27]. Lederberg was unenthusiastic about classical genetics when he arrived at Columbia College in 1941. His interest in "understand[ing] the chemical nature of life" led him to spend much of the next four years studying chemistry, cytology and physiological embryology. But he was not ignorant of classical genetics and the Columbia biologists were well connected with the New York network of scientific communication about genetics. Dobzhansky was a central figure. Arthur Pollister was in close touch with researchers at the Rockefeller Institute. Alfred Mirsky worked at both institutions. Lederberg not only learned quickly about the neo-darwinian developments described earlier but he also heard about the work of Avery *et al* from Mirsky and promptly read their paper. If the Avery *et al* work sharpened Lederberg's interest in bacterial reproduction, Dubos' review of the inconclusiveness of evidence on sexual reproduction sharpened his scepticism, the cognitive and structural elements were coalescing.

In Lederberg's second year at Columbia he met Francis Ryan, an assistant professor, who had just completed a postdoctoral fellowship at Stanford with Beadle and Tatum. It was Ryan who first told him about the work on biochemical genetics and who persuaded him that chemistry and genetics were not as far apart as he had thought[27]. It was also Ryan who generously provided Lederberg with laboratory facilities, catalysed his association with Tatum, and, most importantly, encouraged, educated and socialized him as a scientist. Columbia provided Lederberg with a multifaceted and advantageous structural context for his scientific development and for the initiation of a high-risk, high-stakes research programme.

Lederberg's plan for research was well worked out by July 1945, when he was a second-year student at Columbia Medical School but continued to work in Ryan's laboratory. The research might have been pursued at Columbia, but Ryan and Lederberg both recognized that an association with Tatum would be valuable. In particular, his experience in microbial biochemistry could help broaden Lederberg's education beyond the opportunities available on Morningside Heights. Furthermore, Tatum, then in the process of moving to Yale, was rapidly becoming recognized as a scientific leader. He could provide Lederberg with better access not only to information, research materials and fellowship support, but also to the invisible college of the emerging scientific discipline of biochemical genetics. The impact of such informal ties between investigators on the directions and pace of scientific research has yet to be properly investigated.

Lederberg's status as a medical student was less an obstacle to his investigating bacterial sex than might be supposed. Though much of his time was spent on course work, he was not subject to the constraints that apply in the early years of study for the PhD He did not, like ordinary graduate students, have to choose a research problem that would be suitable for a thesis and publication Being marginal[40] to the biological research enterprise, he could afford to take on a high-risk problem The search for bacterial sex was definitely high-risk, it was not one likely to produce useful and publishable findings After all, not observing bacterial recombination would scarcely demonstrate that it did not exist The risk of a negative finding using *E coli* is now known precisely, bacterial recombination being observed in only five per cent of all strains with the techniques used in 1946

For a different set of reasons, Tatum could also afford research on a high-risk problem at the time He had a variety of projects in process in his laboratory and could manage to take a long-shot experiment that required little time and little money For both men, bacterial recombination was a good gamble, failure would have low marginal costs for each but promised large if prospectively improbable returns High-risk investigations are not equally feasible for all scientists They fall to the comparatively well-established or to those who are marginal as Lederberg was in 1946 Those who solve high-risk problems, having chosen them in the first place, may more often come from the ranks of the well established than from neophytes thus contributing to he accumulation of advantage[41] Risk-taking in science is a matter not only of psychological daring but also of position in the social structure[14,42]

After a brief correspondence, Tatum invited Lederberg to work at Yale He arrived in March 1946, genetic recombination in *E coli* was experimentally observed early in May The results were so arresting that Tatum arranged for Lederberg to present them at the Cold Spring Harbor Symposium to be held in July The publications which followed[3,4,43] did not merely describe the results of the initial laboratory investigation They are the product of critical discussion of those results at that meeting and of follow-up experiments done immediately afterwards[27] The dynamics or organized scepticism in science[44] can be observed in the records of that meeting and later in responses to the papers announcing the discovery Even as first published, discoveries are not simply reports of events intially observed in the laboratory[45,46] but often are also the outcomes of exchange between contributors and their critics Treating scientific contributions as the results of inquiry, criticism and subsequent work makes problematic the custom of designating this or that scientist as the exclusive contributor and focusses attention on the operation of organized scepticism and its effect on shaping the meaning and assessment of those contributions

Conclusions

Was the investigation of sexual recombination in bacteria postmature, that is, conducted significantly later than it could have been? The problem was obscured for decades by the Cohn–Koch dogma of monomorphism and the conviction that bacterial variation resulted only from contamination This was so even for Beijerinck and members of the Delft School who did not subscribe to strict monomorphism, knew how to mark microbial strains by their fermentative and nutritional characteristics (the basis of Lederberg's design), knew about Mendelian segregation in plants, and might have appreciated the significance of sexual recombination in bacteria were it observed But, they were committed to the view that bacteria reproduced only by fission and did not consider the phenomenon problematic In principle, the investigation was technically feasible by 1908, as demonstrated by Browning's[47] use of drug resistance as a selective marker, an early anticipation of the Lederberg–Tatum work But Browning dealt with a different organism, reported a negative result, and used terminology not readily transferable to the case of bacterial recombination In the 1930s, bacterial sex was still viewed as unlikely, even as a disputable idea Yet had it been demonstrated experimentally, it would have been understood and appreciated by geneticists and possibly even by bacteriologists

This case study suggests that problem-identification and selection in science have features deserving further analysis First, the solutions to two classes of problems are apt to be postmature those which do not survive competition for scientists' attention when they first appear because they seem insignificant, unfeasible or both and those which are obscured by prevailing cognitive commitments or have no socially and cognitively defined disciplinary home Second, in calculating the probable returns on selecting problems for investigation, scientists assess the likelihood of error and this contributes to the continuing neglect of certain problems that have a history of being error-prone Third, the feasibility of addressing high-risk problems in science and so of making major advances in this way is not equal for all investigators, they are left largely to the well-established who can afford them and to others who have a smaller stake, for structural reasons, in their immediate record of publication What scientists define as problematic and worthy of investigation are the products of interactions between cognitive and social processes

Research supported by NSF Foundation (SES 80-08609), the Russell Sage Foundation and the Center for Advanced Study in the Behavioural Sciences

---

Harriet Zuckerman is at the Department of Sociology, Columbia University, New York, New York 10027, USA and Joshua Lederberg is at The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

---

1 Barber, B *Science* **134**, 596 – 602 (1961)
2 Stent, G S *Sci Am* **227**, 12, 84 – 93 (1972)
3 Lederberg, J & Tatum, E L *Nature* **158**, 558 (1946)
4 Lederberg, J & Tatum, E L *Cold Sp Harb Symp* **9**, 113 – 114 (1946)
5 Nevins, A *Wilson Lib Bull* **60**, 607 – 615 (1966)
6 Bension, S in *Modern Methods in the History of Medicine* (ed Clarke, E) 286–365 (Athlone, London, 1971)
7 Starton, G *The Study of the History of Mathematics* **35** (Harvard University Press, Cambridge, 1936)
8 Merton, R K *The Sociology of Science* 371 – 382 (University of Chicago, 1973)
9 Garfield, E *Essays of an Information Scientist (1979 – 80)* 488 – 493 (ISI, Philadelphia, 1980)
10 Cole, S *Am J Sociol* **76**, 286 – 306 (1970)
11 Garfield, E *Current Comments* **1717**, 3 – 10 (1985)
12 Butterfield, H *The Whig Interpretation of History* (W W Norton, New York, 1965)
13 Pauling, L *Nature* **248**, 769 – 771 (1974)
14 Zuckerman, H *Sociol Inq* **48**, 65 – 95 (1979)
15 Holton, G J *Thematic Origins of Scientific Thought Kepler to Einstein* (Harvard University Press, Cambridge 1973)
16 Weinberg, S *Daedalus* **106** 17 – 35 (1977)
17 Geison, G L in *Dict Sci Biog* **3** (ed Gillispie C C) 335 – 41 (Scribner's, New York, 1971)
18 Dubos, R J *The Bacterial Cell* **135** (Harvard University Press, Cambridge, 1945)
19 Merton, R K *Social Theory and Social Structure* 475 – 490 (Free Press, New York, 1968)
20 Zuckerman, H in *Deviance and Social Change* (ed Sagarin, E A) 87–138 (Sage Publishing Beverly Hills California, 1977)
21 Ben-David, J in *Culture and Its Creations* (eds Ben-David J & Clark, T) 255–65 (University of Chicago Press Chicago, 1977)
22 *Kluyver, Albert Jan His Life and Work* (eds Kamp A R , La Rivière J W M & Verhoeven W) **186** (Amsterdam North Holland 1959)
23 Beijerinck, M *Versl Akad Wetensnch (Amsterdam)* **9** 310 (1901)
24 Aronson, N *Why Weren t Vitamins Discovered Earlier?* (unpublished, 1984)
25 Dobzhansky, T G *Genetics and the Origin of Species* (Columbia University Press, New York 1937)
26 Beadle, G W & Tatum E L *Proc natn Acad U S A* **27** 499 – 506 (1941)
27 Lederberg, J *Nature* **324**, 627-631 (1986)
28 Lederberg J & Tatum, E L *A Rev Genet* **13**, 1–5 (1979)
29 Oparin A I *The Origin of Life on Earth* Trans Ann Synge (Academic, New York, 1957)
30 Schrodinger, E *What is Life? The Physical Aspect of the Living Cell* (Cambridge University Press Cambridge 1962)
31 Avery, O T MacLeod, C M & McCarty M *J exp Med* **79** 596 – 602 (1944)
32 Judson, H F *The Eighth Day of Creation Makers of the Revolution in Biology* (Simon & Schuster New York 1979)
33 McCarty M *The Transforming Principle Discoveries That Genes Are Made of DNA* (Norton New York 1986)
34 Olby R C *The Path to the Double Helix* (University of Washington Press Seattle 1974)
35 Dubos R J *The Professor, The Institute and DNA* (Rockefeller University Press New York 1976)
36 Fruton J S *Molecules and Life Historical Essays on the Interplay of Chemistry and Biology* (Wiley New York 1972)
37 Lederberg J *Nature* **239**, 234-236 (1972)
38 Cohen J S Portugal F H *Persp Biol Med* **18** 204 – 207 (1975)
39 Dubos R J *The Bacterial Cell* **181** (Harvard University Press Cambridge 1945)
40 Gieryn T F & Hirsch R F *Soc St of Sci* **13** 87 – 106 (1983)
41 Merton R K *The Sociology of Science* 439 – 59 (University of Chicago Press Chicago 1973)
42 Mulkay M J *The Social Process of Scientific Innovation A Study in the Sociology of Science*, 49 – 51 (Macmillan London 1972)
43 Tatum E L & Lederberg J *J Bact* **53** 673 – 684 (1947)
44 Merton R K *The Sociology of Science* 267 – 278 (University of Chicago Press Chicago 1973)
45 Latour B & Woolgar S *Laboratory Life The Social Construction of Scientific Facts* (Sage Publications Beverly Hills 1979)
46 Knorr – Cetina K *The Manufacture of Knowledge An Essay on the Constructivist & Contextual Nature od Science* (Pergamon Oxford 1981)
47 Browning C H *J Path Bact* **12**, 166 – 190 (1908)

Helium isotopes in sedimentary basins

E. R. Oxburgh, R. K. O'Nions & R. I. Hill*

Department of Earth Sciences, University of Cambridge, Downing St, Cambridge CB2 3EQ, UK

Measurements of helium isotopes in the ground waters and natural gases of sedimentary basins reveal that some possess a major component of mantle-derived helium, apparently related to their mechanism of formation. Neogene basins formed by loading do not involve introduction of significant mantle helium, whereas those formed by extension do.

MANTLE degassing occurs principally through young oceanic lithosphere, and to a lesser extent through tectonically active portions of continental crust[1-3]. The phenomenon is particularly well documented for mantle-derived helium, which contains more ^{3}He than crustal helium of purely radiogenic derivation. Stable parts of the continents are essentially impermeable to mantle volatiles, and helium from even the deepest boreholes has an isotopic composition characteristic of radiogenic production. But the formation of sedimentary basins involves deformation of the continental lithosphere which, depending on the processes involved, may or may not allow transient transfer of mantle-derived fluids into the crust. Here we suggest that the isotopic composition of helium in the ground waters and natural gases sampled from sedimentary basins may indicate the processes that acted during their formation.

Sedimentary basins may develop in different tectonic settings, and often the mechanisms involved in their formation remain poorly understood. Studies of the rates of basin subsidence, and the spatial and temporal variations of these rates, have shown that some basins have formed largely through thermal subsidence, whereas others appear to result from tectonic loading of the crust. Various explanations have been offered for basin formation through thermal subsidence[4-7]: one involves extension (or boudinage) of the lithosphere and upwelling of asthenosphere within the stretched zone[6], another involves stretching of the lithosphere without thinning, extension being accommodated by the emplacement of mantle-derived dykes[7]. These are not mutually exclusive alternatives, and elements of both could be present in the development of any one basin. The North Sea basin provides a good example of thermal subsidence[8,9], and the demonstration of crustal thinning as well as aspects of the early subsidence history have suggested that this basin developed largely through lithospheric stretching.

When lithospheric stretching is accompanied by passive uprise of the asthenosphere, decompression may produce some melting[10], and mantle volatiles may be released into the crust. When coupled with the elevated temperatures expected in the thinned crust[6], such processes may facilitate the transport of mantle helium to shallow crustal levels. The dyke emplacement model for lithospheric extension[7] could involve the release of mantle-derived volatiles from basaltic melts emplaced into the crust beneath the developing sedimentary basin.

Loading basins form where the crust is depressed adjacent to a tectonically imposed load, such as a stack of thrust sheets. Mantle melting is not expected to accompany the development of such basins, and enhanced transfer of mantle-derived fluids into the crust is not anticipated. The Molasse and East Carpathian basins of the Alpine chain and many of the basins east of the Canadian Rockies are examples of loading basins.

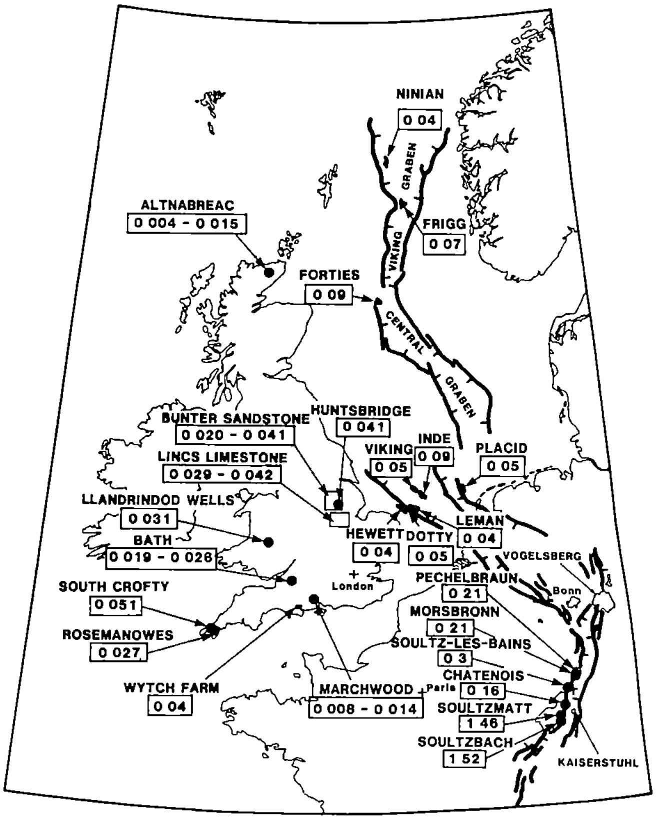


Fig. 1 Isotopic composition of helium in ground waters and natural gases in the mainland United Kingdom, North Sea and Rhinegraben. The values shown are for R/R_a, the ratio of the ^{3}He/^{4}He ratio of the sample (R) to that of the atmosphere (R_a). Natural gas samples from the North Sea basin have R/R_a values close to the expected radiogenic production ratio. A substantial mantle-derived helium component is present in samples from the southern part of the Rhinegraben. Data from ref. 12 and unpublished work.

Helium isotope results for sedimentary basins

Helium isotope data are available for ground-water and natural gas samples from a number of sedimentary basins of contrasting age and style[2,11-13] (Figs 1-3). The Molasse basin north of the Alps, the Pannonian basin of Hungary, the Rhinegraben and the Niigata basin of western Honshu (Japan) are used as examples of young (Neogene) sedimentary basins, the Mesozoic and Palaeogene North Sea basin and the Palaeozoic Hugoton–Panhandle gas fields provide information on older structures. Of these, only the Molasse basin is clearly formed by loading. Samples from the Niigata basin[11], the North Sea[12] and the Hugoton–Panhandle[13] are all from producing gas wells, samples from the Rhinegraben[12] and from the Molasse and Pannonian basins (Fig. 2) were collected from producing water wells, and from deep-source springs. The two sampling modes each provide representative samples of the local crustal helium, in many cases

---

* Present address: Research School of Earth Sciences, Australian National University, Canberra, Australia.

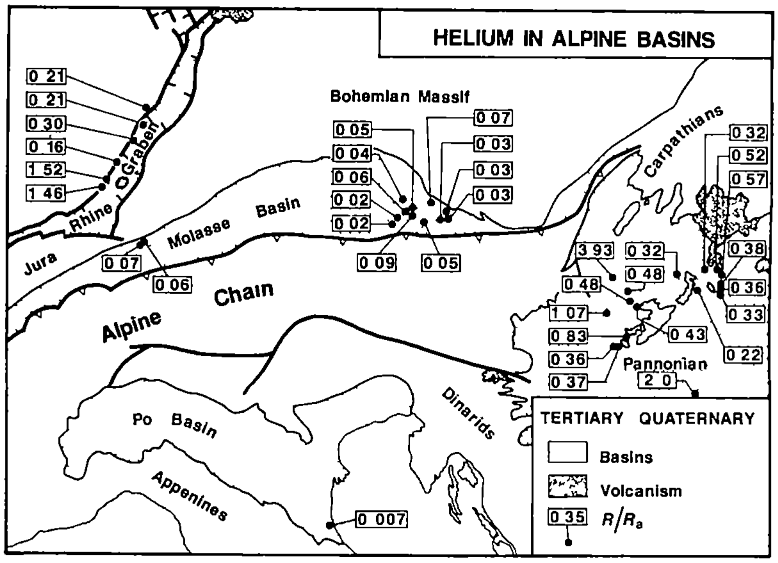


Fig. 2 R/R_a values for ground waters from Neogene basins Values for the Molasse basin (refs 30, 31 and W Balderer, personal communication) show little evidence of the presence of mantle-derived helium In contrast, the Pannonian basin samples[32] all contain a mantle component

it has been demonstrated that natural gases and formation waters from a single basin have comparable ^{3}He/^{4}He ratios[14–16] The ^{3}He/^{4}He ratios (R) for samples from the various basins have been normalized to the atmospheric ^{3}He/^{4}He ratio ($R_a = 1\ 41 \times 10^{-6}$) and are presented in Figs 1 and 2

Figure 3 summarizes the available data, which range from 0 02 R_a to 6 1 R_a Interpretation of the two extremes of the histogram is straightforward values >4 R_a indicate a substantial mantle helium component ($\geq 50\%$), whereas values close to 0 02 R_a indicate such low ^{3}He abundances that no mantle helium component is recognizable The isotopic composition of radiogenic crustal He is controlled essentially by the abundance of lithium ^{4}He is generated by the ^{238,235}U, ^{232}Th decay series, whereas ^{3}He is derived from ^{6}Li by an (n, α) reaction The composition of radiogenic helium may vary locally within a rock, reflecting the distribution of uranium, thorium and lithium between different phases and on different length scales This problem has been examined in some detail[16–19] and it appears that the composition of helium released from rocks is averaged over many possible generation sites As helium samples are representative of relatively large ground-water or natural gas systems, for normal crustal lithologies radiogenic helium R/R_a is expected to average $\leq 0\ 04$, which is about the upper limit for water samples collected from different parts of the stable United Kingdom crust (refs 12, 20 and unpublished data) Although local departures may arise from anomalously lithium-rich lithologies, R/R_a values $>0\ 08$ are interpreted here as being mixtures of radiogenic helium ($R/R_a \approx 0\ 04$) and mantle helium ($6 < R/R_a < 10$) Anthropogenic ^{3}He (derived from bomb tritium) can be present only in very young (post-1950) ground-water components and needs to be considered only in the case of Molasse basin waters The other fluids analysed either have helium contents too high for their composition to have been significantly affected by bomb tritium, or are demonstrably too old to be modified in this way, or both

Neogene basins The Niigata, Molasse and Pannonian basins began to form in the latest Palaeogene or early Neogene The Niigata basin is an epicontiental back-arc basin behind the active northern Honshu volcanic arc Subsidence began at ~23 Myr BP, and a thick sequence of volcanic and volcaniclastic rocks was deposited on a basement of Palaeozoic metamorphic rocks and Cretaceous granites Volcanic input to the basin has waned since the end of the Miocene (6 Myr BP), although clastic sedimentation and deformation have continued to the present day

Unlike the Niigata basin, whose development accompanied the opening of the Japan Sea[21], the Pannonian basin is entirely intra-continental It truncates many of the thrust units of the Alpine belt, and subsidence is associated with extensional and strike-slip faulting, although the broader tectonic setting is not entirely clear[22,23] Volcanism within the Pannonian basin continued until the end of the Pliocene The southern Rhinegraben is a narrow intra-cratonic rift which has been accumulating sediment since the Palaeogene The region is seismically active today[24], and Tertiary volcanism continued until 11 Myr BP The Molasse basin, in contrast, is a classic loading basin formed in the late Palaeogene and Neogene, and is the site of substantial Miocene deposition The basin owes its origin to northward transport of the Upper East Alpine nappes and consequent loading of the Alpine foreland

R/R_a values for helium from the Pannonian basin range from 0 22 to 3 9 and indicate the presence of a mantle-derived component ranging from 2 5 to 50% The values for the Rhinegraben are similar but a little lower, (0 21–1 3)R_a Those from the Niigata basin, however, are in some cases as high as 6 1 R_a, indicating that locally the crust is completely swamped with mantle-derived helium The Molasse basin helium has R/R_a values of 0 02–0 09, and nearly all lie within the anticipated range for radiogenic helium generated within the crust The two highest values ($>0\ 07$) have the largest air corrections and the presence of anthropogenic ^{3}He is suspected, if this explanation of the high values is incorrect, the maximum amount of mantle helium in the Molasse basin samples is 0 9% The Molasse basin shows no evidence for either volcanic activity contemporaneous with basin development nor evidence for crustal thinning For these four young basins, the presence or absence of mantle helium correlates with tectonic setting

Older basins. We now consider two older basins The development of the North Sea basin began with two phases of rifting, in the late Carboniferous (~300 Myr) and Triassic (210–250 Myr), followed by Mesozoic deposition in fault-controlled extensional troughs During the Tertiary and Quaternary there was broad, regional subsidence, with the centre of the basin subsiding more than the margins R/R_a values for helium from major North Sea gas fields and from the adjacent mainland United Kingdom ground waters (Fig 1) range from 0 04 to 0 09 Contamination of these samples by anthropogenic ^{3}He is unlikely Of the three samples with $R/R_a > 0\ 07$, two are from fields adjacent to bounding faults of the Central and Viking grabens These relationships have been discussed in detail by Hooker *et al*[12], who pointed out that the North Sea gas reservoirs are dominated by crustally produced radiogenic helium The weak mantle signature which is locally present could have been acquired either at the time of hydrocarbon migration, or some time later by the leakage of mantle-derived helium upwards along major faults

The Hugoton–Panhandle region of the south-central United

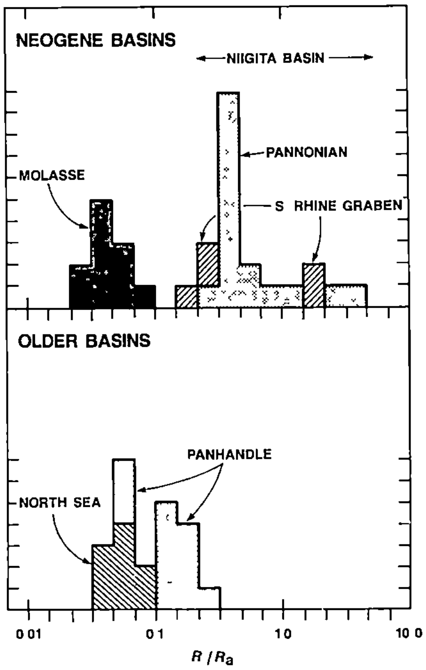


Fig 3 Histograms of helium isotope compositions in Neogene and older sedimentary basins Data sources as for Figs 1 and 2, and from ref 11 for the Niigata basin Note that the plotting intervals on the horizontal axis are unequal because the data are presented on a logarithmic scale

States is of particular interest because it is one of the largest accumulations of natural gas discovered, and has helium contents of up to 2% The gas reservoirs are in rocks of uncertain tectonic setting The southern 'Panhandle' part of the reservoir is draped over a late Precambrian and early Palaeozoic basement high, the Amarillo uplift, and the northern 'Hugoton' fields are situated in sediments of the western margin of the Palaeozoic Anadarko Basin The early history of the region appears to involve an early and rapid subsidence (from 525 to 500 Myr) with little accompanying volcanism, followed by an extended period of sedimentation until ~435 Myr (refs 25, 26)

R/R_a values for Panhandle gases vary from 0 05 to 0 21, and all but one of the thirteen samples analysed by Hill *et al*[13] show a clear mantle-derived component which reaches a maximum of ~2% It is not known when the ^{3}He signature was acquired, and the helium is not necessarily residual from the time of basin formation[13], but could be a recent or contemporary feature associated with igneous activity beneath the northern part of the field

Discussion

Present knowledge of the processes that control the release of mantle volatiles and their ascent through the crust is extremely rudimentary It is clear that igneous activity is very effective in this way, and almost everywhere that measurements have been made on gases issuing from fumaroles, volcanic vents or hot springs associated with volcanic activity, the presence of mantle helium has been detected The three areas described above that have strongest mantle signatures, the Niigata and Pannonian basins and the Rhinegraben, have been the loci of Palaeogene and/or Neogene igneous activity and it could be suggested that degassing of helium from the mantle occurs only in association with melting and therefore with igenous activity We now consider this possiblity further, in the context of sedimentary basin formation

For thermal subsidence models of sedimentary basin formation, the introduction of mantle melts into the crust during the extensional phase is an explicit feature of the dyke injection model and a possible consequence of asthenospheric decompression in stretching models Mantle helium would accompany melts into the crust, and could reach the upper crust before the maturation and migration of hydrocarbons

Of the extensional basins that we have considered, three, the Pannonian basin, the Rhinegraben and the Niigata basin, have been active in the Neogene and are active to varying extents today All have mantle helium in their ground waters, which presumably reflects degassing of helium recently derived from the mantle in each of these areas Ground-water helium is likely to be lost to the atmosphere relatively quickly, and once degassing ends there may be no indication of the former presence of mantle volatiles in these areas, unless samples of crustal fluids are trapped and preserved in some way Torgersen and Clarke[27] have recently argued, from helium concentrations of in ground waters, that degassing of helium from the continental crust is essentially a steady-state process

The North Sea is essentally a Mesozoic basin in which slow subsidence in the axial zone continues today, whereas the Panhandle basin is of Palaeozoic age and has been inactive during the latter part of the Mesozoic and the Tertiary The evidence for mantle volatiles in both cases comes from gases trapped in hydrocarbon reservoirs The question in both cases is when was the mantle gas introduced? If the release of mantle gases associated with basin formation ends before the development of traps, no gases are likely to be preserved The persistence of a mantle helium signature in a sedimentary basin after its formation clearly depends on how long the mantle supply continues, and on the pathways followed by the mantle fluid in the crust At the two extremes will be the gas trapped in the 'tight' reservoir, with the possibility of long retention, and the gas entrained in ground water that makes its way rapidly to the surface to be lost to the atmosphere The decay times associated with these two extremes could range from several hundreds to less than tens of millions of years In the former case the signature will also weaken with time through dilution with locally produced radiogenic crustal helium

In the North Sea, hydrocarbon migration seems in some cases to have occurred nearly 150 Myr after the main extensional phase of basin development If the extensional phase was accompanied by emplacement of mantle-derived melts, then the general absence of mantle ^{3}He in the basin today would require that helium be effectively lost in $\leqslant$100 Myr The presence of mantle ^{3}He in the much older Hugoton–Panhandle rocks must imply either long-term retention from the Palaeozoic and therefore a quite different effective diffusivity structure for the crust compared with that of the North Sea, or alternatively a more recent input of mantle-derived gas

In the case of the North Sea we note that, away from the Central Graben, all well gases analysed from the basin, and all groundwaters from its landward extension, show only radiogenic crustal helium The only sign of mantle helium is a weak signature found along the distant northward extension of the active Rhinegraben faulting system into the North Sea, and it is tempting to suppose that there was a small amount of a localized mantle gas leakage at its northern end

It is not clear whether the ^{3}He-rich helium is derived from the asthenosphere or the lithosphere The convecting upper mantle, as sampled at spreading-ridge axes, is a well-mixed reservoir for helium with $6 < R/R_a < 10$, but the continental lithospheric mantle is also known to contain helium Xenoliths of continental lithospheric mantle material, brought to the surface by recent volcanic activity, contain helium with a range of R/R_a values similar to those characterizing mid-ocean-ridge basalts (MORB)[28], however, it is not known whether helium of such a composition is to be found everywhere in continental lithospheric mantle Helium in fluids from a wide range of circum-Pacific volcanoes has a relatively restricted compositional range around $R/R_a = 6$, and is interpreted as being a mixture of MORB-type helium from mantle beneath the volcanic arc and radiogenic helium derived from the subducted crust[29]

In the absence of evidence to the contrary, we assume that mantle helium is transported into the crust in melts, from which it subsequently degasses, and that melting occurs principally in the hotter asthenospheric material which undergoes decompression beneath at least some sedimentary basins formed by thermal subsidence Note, however, that non-magmatic fluids of mantle origin could satisfy the helium observations equally well

In summary, it appears that the presence or absence of mantle-derived ^{3}He at shallow levels in Neogene sedimentary basins relates to their mechanism of formation The presence of significant amounts of mantle helium in the Pannonian and Niigata basins and the Rhinegraben is related to the introduction and degassing of mantle-derived fluids produced as an integral part of the extensional process Extension by dyke injection cannot be easily distinguished from extension by thinning, although some resolution may ultimately be possible because of predicted differences in the amounts of melt introduced The Molasse basin, a classic loading basin, has developed without the introduction of identified mantle-derived melts or helium The interpretation of the helium isotope composition in sedimentary basins whose development began in the Palaeozoic or Mesozoic, such as the Hugoton-Panhandle and North Sea basins, is less straightforward and requires some knowledge of the time required for the loss of helium in crustal regions undergoing basin development Both of these basins have probably received late additions of mantle volatiles Parts of this work were carried out under a research contract with the European Community, no EN3G 004 UK(H) The ideas presented in this paper have been discussed with many colleagues, for whose comments we are extremely grateful

Department of Earth Sciences contribution ES810

Received 21 July, accepted 24 October 1986

1 Craig, H & Lupton, J E *Earth planet. Sci Lett* **31**, 369-385 (1976)
2 Polyak, B G, Tolstikhin, I N & Yakutseni, V P *Geotectonics* **13**, 339-351 (1979)
3 Mamyrin, B A & Tolstikhin, I N *Helium Isotopes in Nature* (Elsevier, Amsterdam, 1984)
4 Sleep, N H *Geophys J R. astr Soc.* **24**, 325-350 (1971)
5 Falvey, D A. *APEA J* **14**, 95-106 (1974)
6 McKenzie, D *Earth planet Sci Lett* **40**, 25-32 (1978)
7 Royden, L, Sclater, J G & von Herzen, R. P *Bull. Am. Ass Petrol Geol.* **64**, 173-187 (1980)
8 Sclater, J G & Christie, P A F *J geophys. Res* **85**, 3711-3739 (1980)
9 Wood, R. & Barton, P *Nature* **302**, 134-136 (1983)
10 McKenzie, D P *J Petrol.* **25**, 713-765 (1984)
11 Wakita, H & Sano, Y *Nature* **305**, 792-794 (1983)
12 Hooker, P J, O'Nions, R. K & Oxburgh, E R *Nature* **318**, 273-275 (1985)
13 Hill, R. I, O'Nions, R K, Oxburgh, E R & Gold, T (in preparation)
14 Kamensky, I L *et al. Geokhimiya* **5**, 682-695 (1976)
15 Matveyeva, E S, Tolstikhin, I N & Yakutseny, V P *Geokhimiya* **3**, 307-317 (1978)
16 Polyak, R. G, Kononov, V I, Tolstikhin, I N, Mamyrin, B A & Khabarin, L V *Therm chem. Problems hydrol. Sci. Publ.* **119**, 17-33 (1976)
17 Gerling, E K., Mamyrin, B A, Tolstikhin, I N & Yakovleva, S S *Geokhimiya* **5**, 608-617 (1971)
18 Morrison, P & Pine, J *Ann. N Y Acad. Sci.* **62**, 69-92 (1955)
19 Andrews, J N *Chem. Geol.* **49**, 339-351 (1985)
20 Hilton, D R., Oxburgh, E R. & O'Nions, R. K. in *High Heat Production (HHP) Granites, Hydrothermal Circulation and Ore Genesis*, 135-142 (Institute of Mining and Metallurgy, London, 1985)
21 Katahira, T & Ukai, M *Mem. Am. Ass Petrol. Geol.* **25**, 276-279 (1976)
22 Royden, L H, Horvath, F & Burchfiel, B C *Geol. Soc. Am Bull.* **93**, 717-725 (1982)
23 Sclater, J G *et al. Earth planet Sci. Lett.* **51**, 139-162 (1980)
24 Illies, J H *Tectonophysics* **73**, 249-266 (1981)
25 Feinstein, S *Bull. Am. Ass Petrol. Geol.* **65**, 2521-2533 (1981)
26 Garner, D L & Turcotte, D L *Tectonophysics* **107**, 1-24 (1984)
27 Torgeser, T & Clarke, W B *Geochim. cosmochim. Acta* **49**, 1211-1218 (1985)
28 Porcelli, D R, O'Nions, R. K. & O'Reilly, S Y *Chem. Geol.* **54**, 237-250 (1986)
29 Lupton, J E *A. Rev Earth planet Sci* **11**, 371-414 (1983)
30 Andrews J N *et al. Earth planet Sci. Lett.* **73**, 317-332 (1985)
31 Balderer, W *Nagra tech. Rep* No NTB 85 06 (Baden, 1985)
32 Deak, J *et al.* in *A Case Study on the Pannonian Basin* (eds Horvath, F & Royden, L) (*Am. Ass Petrol. Geol., Spec. Pub* in the press)

The c-*erb-A* protein is a high-affinity receptor for thyroid hormone

Jan Sap[*], Alberto Muñoz[*], Klaus Damm[*], Yves Goldberg[†], Jacques Ghysdael[†], Achim Leutz[*‡], Hartmut Beug[*] & Björn Vennström[*]

[*] Differentiation Programme, European Molecular Biology Laboratory, Postfach 10 2209, D-6900 Heidelberg, FRG
[†] INSERM, Pasteur Institute, 15 Rue Guerin, 59019 Lille Cedex, France

*Hormone binding and localization of the c-*erb-A *protein suggest that it is a receptor for thyroid hormone, a nuclear protein that binds to DNA and activates transcription In contrast, the product of the viral oncogene v-*erb-A *is defective in binding the hormone but is still located in the nucleus*

THE avian v-*erb-A* oncogene does not by itself cause tumours in animals, but enhances the erythroblast-transforming potential of primary oncogenes, such as v-*erb-B* or other sarcoma-inducing oncogenes[1,2] Two specific effects of v-*erb-A* have been characterized First, erythroblasts transformed by the primary oncogene alone spontaneously differentiate into erythrocytes, and v-*erb-A* completely inhibits this maturation Second, v-*erb-A* allows transformed erythroblasts to propagate in media with wide ranges in *p*H and salt concentration[1-3] These observations suggest that v-*erb-A* in erythroid cells affects ion transport systems and regulatory factors of differentiation[3]

It was recently demonstrated that the v-*erb-A*-specific part of the $p75^{gag\text{-}v\text{-}erb\text{-}A}$ protein (encoded by the avian erythroblastosis virus, AEV-ES4) bears homology to receptors for steroid hormones[4-7], suggesting that the effects of v-*erb-A* could be due to direct transcriptional deregulation of genes responsible for the aforementioned effects[3] To understand these processes better we have isolated and characterized complementary DNA clones of the cellular *erb-A* gene as well as their proteins, and demonstrate that the c-*erb-A* protein has biochemical properties very similar to those of the receptor for the thyroid hormones triiodothyronine (T_3) and thyroxine (T_4) (see ref 8 for review)

Cloning and sequencing of cDNAs for c-*erb-A*

The avian c-*erb-A* gene is transcribed into two messenger RNAs, 4 5 and 3 kilobases (kb) long, that are present at low levels (5-10 copies per cell) in many adult and embryonic tissues (ref 9 and B V, unpublished data) We therefore screened with a v-*erb-A* probe a chicken embryonic cDNA library prepared in λ bacteriophage gt11 (refs 10, 11) Ten positive clones were isolated from the 10^6 plaques screened, and two (clones F1 and U1), that hybridized to probes representing 5′ and 3′ domains in v-*erb-A*, were subjected to nucleotide sequence analysis Figure 1 shows the presence of a long open reading frame of 1,224 nucleotides homologous to v-*erb-A*, suggesting a relative

---

[‡] Present address Department of Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11790, USA

```
                                                                                 -288 CTCGGCCCCGCTCCGCTCCGCTCCTCTCCGCGCTCCGCCCCGGCCCCG   -240
              CTCCGCGCCCCGCACGGCCCCGTTCTCAGCAGGAGTGCTCGCGGAGGAGGCAGCGCCCAGCAGCCCCCAGCGCCCCCCACCTCTGCCCCCCCCCCCGGGATGCTGCAGGGAGCCCAGCAC   -120
              TGAGGTGGGCGCCATGGGGCGCGCATCGTAGCCCCCCCCCGCCGTGGCCCCCCCGCGCGTGGCACCGGTGGCGGAGCCCACTCGGCCGGTGCTGGGCGCTCCTTGATGGAATTGCGGTGA   -1
                                            . gag | v-erbA
v-erbA                           |CCTGGACCC-CTCT-GGAGCCA|                           A
c-erbA     1  ATGGAACAGAAGCCCAGCACCCTGGACCCGCTGTCGGAGCCAGAGGACACTCGGTGGCTGGATGGCAAGCGCAAAAGAAAGAGCAGCCAATGTTTGGTGAAGAGCAGCATGTCAGGGTAC   120
c-erbA        MetGluGlnLysProSerThrLeuAspProLeuSerGluProGluAspThrArgTrpLeuAspGlyLysArgLysArgLysSerSerGlnCysLeuValLysSerSerMetSerGlyTyr
v-erbA                                                                             His

                        G                              T                                                      A
         120  ATCCCTAGCTACCTGGACAAAGATGAACAGTGCGTGGTGTGCGGGGACAAAGCCACCGGCTACCACTACCGCTGCATCACCTGCGAGGGCTGCAAGGGCTTTTTCCGTCGGACCATCCAG   240
              IleProSerTyrLeuAspLysAspGluGlnCysValValCysGlyAspLysAlaThrGlyTyrHisTyrArgCysIleThrCysGluGlyCysLysGlyPhePheArgArgThrIleGln
                                                                                                              Ser
                       Cys                   --->

                                          C
         241  AAGAACCTGCACCCCACCTACTCCTGCAAGTACGATGGGTGCTGCGTCATCGACAAGATCACCCGCAACCAGTGCCAGCTGTGCCGCTTCAAGAAGTGCATCTCCGTGGGCATGGCCATG   360
              LysAsnLeuHisProThrTyrSerCysLysTyrAspGlyCysCysValIleAspLysIleThrArgAsnGlnCysGlnLeuCysArgPheLysLysCysIleSerValGlyMetAlaMet
                                         Thr                                                                             <---
                                                                                                       •
                                                                                  A
         361  GACCTGGTGCTGGACGACTCGAAGCGCGTAGCCAAGCGGAAGCTGATCGAGGAGAACCGGGAGCGGCGGCGCAAGGAGGAGATGATCAAATCCCTGCAGCACCGGCCCAGCCCCAGCGCA   480
              AspLeuValLeuAspAspSerLysArgValAlaLysArgLysLeuIleGluGluAsnArgGluArgArgArgLysGluGluMetIleLysSerLeuGlnHisArgProSerProSerAla

                                                                                          G              T
         481  GAGGAGTGGGAGCTGATCCACGTGGTGACCGAGGCGCACCGCAGCACCAACGCGCAGGGCAGCCACTGGAAGCAGAAGAGGAAATTCCTGCCCGAAGATATCGGTCAGTCGCCCATGGCC   600
              GluGluTrpGluLeuIleHisValValThrGluAlaHisArgSerThrAsnAlaGlnGlySerHisTrpLysGlnLysArgLysPheLeuProGluAspIleGlyGlnSerProMetAla
                                                                                         Arg            Leu

                     T                                                                                          C
         601  TCCATGCCTGACGGGGACAAAGTGGACCTGGAGGCGTTCAGCGAGTTTACAAAAATCATCACGCCGGCCATCACCCGCGTGGTCGACTTTGCCAAAAAACTGCCCATGTTCTCGGAGCTG   720
              SerMetProAspGlyAspLysValAspLeuGluAlaPheSerGluPheThrLysIleIleThrProAlaIleThrArgValValAspPheAlaLysLysLeuProMetPheSerGluLeu
                    Leu                                                                                       Asn

         721  CCGTGCGAGGATCAGATCATCCTGCTGAAGGGCTGCTGCATGGAGATCATGTCGCTGCGCGCCGCCGTGCGCTACGACCCCGAGAGCGAAACGCTGACGCTGAGCGGGGAAATGGCCGTC   840
              ProCysGluAspGlnIleIleLeuLeuLysGlyCysCysMetGluIleMetSerLeuArgAlaAlaValArgTyrAspProGluSerGluThrLeuThrLeuSerGlyGluMetAlaVal

                                                                         •
                          T          A
         841  AAACGCGAGCAGCTGAAGAACGGGGGGCTGGGGGTCGTGTCTGATGCCATCTTCGACCTCGGCAAGTCGCTGTCTGCCTTCAACCTGGACGACACCGAGGTGGCCCTGCTGCAGGCCGTG   960
              LysArgGluGlnLeuLysAsnGlyGlyLeuGlyValValSerAspAlaIlePheAspLeuGlyLysSerLeuSerAlaPheAsnLeuAspAspThrGluValAlaLeuLeuGlnAlaVal

                                                                             T
         961  CTGCTCATGTCCTCAGACCGGACGGGGCTGATCTGCGTGGATAAGATAGAGAAGTGCCAGGAGACGTACCTGCTGGCGTTCGAGCACTACATCAACTACCGCAAACACAACATTCCCCAC   1080
              LeuLeuMetSerSerAspArgThrGlyLeuIleCysValAspLysIleGluLysCysGlnGluThrTyrLeuLeuAlaPheGluHisTyrIleAsnTyrArgLysHisAsnIleProHis
                                                                             Ser

                    T                    G                        A                                                  C    ------------
        1081  TTCTGGCCCAAGCTGCTGATGAAGGTGACGGACCTGCGCATGATCGGCGCCTGCCACGCCAGCCGCTTCCTGCACATGAAGGTGGAGTGCCCCACCGAGCTCTTCCCCCCCCTCTTCCTC   1200
              PheTrpProLysLeuLeuMetLysValThrAspLeuArgMetIleGlyAlaCysHisAlaSerArgPheLeuHisMetLysValGluCysProThrGluLeuPheProProLeuPheLeu
                    Ser                  Ala                     Tyr                                                Ser    -   -  -  -

              ---------------            A    TACAGATTTTGGGAAGTA
        1201  GAGGTCTTCGAGGACCAGGAGGTGTAGGGCCCCCCCCCCACCGCCACGCG  1250
              GluValPheGluAspGlnGluVal***
               -  -  -  -  -          ***
```

Fig 1 Nucleotide and predicted amino-acid sequence of the c-*erb-A* cDNA clone F1, and comparison with v-*erb-A* Only those nucleotides and amino acids in v-*erb-A* that differ from c-*erb-A* are shown, whereas deletions are indicated with bars The three short open reading frames in the 5′ untranslated region are shown in bold The *gag* sequence homologous to c-*erb-A* and which bridges the *gag*–v-*erb-A* junction is boxed, nucleotide differences in this sequence are shown in bold letters and deletions with bars The region homologous to the DNA-binding domain of steroid receptors is indicated by arrows The U1 cDNA clone starts at nucleotide −84

Methods Polyadenylated RNA was isolated from a pool of 10-day-old chicken embryos[9], and a cDNA library was constructed in λ gt11 essentially as described[10 11] The library (~10^6 plaques) was screened with a nick-translated 2 5-kb *Pvu*II fragment encoding the viral *erb-A* and *erb-B* sequences from AEV Ten c-*erb-A* positive clones were identified, which were further characterized by hybridization with probes specific for the 5′ and 3′ parts of v-*erb-A*. The inserts of two clones (F1 and U1) hybridizing to these probes were subsequently subcloned into the pTZ vectors[14] for nucleotide sequence analysis using the dideoxy method[50] The F1 clone (1 55 kb) was sequenced in its entirety on both strands, whereas the U1 (2 1 kb) clone was subjected to partial analysis No differences between the clones were detected in the coding or 5′ non-coding sequences

molecular mass ~46,000 (M_r 46K) for the c-*erb-A* protein The coding sequence is preceded by a 5′ untranslated region of at least 288 nucleotides, which contains three short open reading frames (shown in bold letters) 5′ to the long open reading frame In addition, a fourth reading frame, lacking an AUG codon, is open for 413 nucleotides (position −287 to 126) Further analyses are required to determine if this reading frame is used for translation by any of the c-*erb-A* mRNAs Interestingly, the 5′ untranslated sequence is very G+C rich, and a computer-assisted analysis suggested that it could form stable secondary structures (data not shown) Several of the cDNA clones isolated contained long (1–2 kb) sequences downstream of the open reading frame, their nucleotide sequences were not investigated in detail as a preliminary analysis revealed no long open reading frames, thus indicating that they represented 3′ untranslated sequences These sequences were also heterogeneous in length, suggesting that most of the cDNA clones were generated by non-specific and internal priming of the first strand cDNA synthesis

A comparison of the coding sequences of the cellular and viral[3,12] *erb-A* genes revealed that v-*erb-A* has undergone 17 point mutations, of which two led to amino-acid substitutions located in the putative DNA-binding domain and 11 in other regions (Fig 1) Furthermore, the v-*erb-A* protein has lost nine amino acids (nucleotides 1,188–1,215) located three amino acids before the carboxy-terminus Finally, a comparison between the c-*erb-A* coding sequence and the *gag* sequences of Rous sarcoma virus[13] or AEV (ref 3) demonstrated a region with 18/22 and 19/22 nucleotide homology, respectively (boxed in Fig 1), which bridges the junction between the *gag* and v-*erb-A* domains in $p75^{gag\text{-}v\text{-}erb\text{-}A}$ As a consequence, the amino-terminus of the c-*erb-A* polypeptide is 12 amino acids longer than the *erb-A*-specific part of $p75^{gag\text{-}v\text{-}erb\text{-}A}$, when taking into account the region homologous to *gag* This also suggests that c-*erb-A* was fused to *gag* either by homologous recombination at the DNA level or during retrotranscription of c-*erb-A* mRNA packaged into virions

Characterization of the c-*erb-A* protein

To determine whether the cDNA clones encode an authentic c-*erb-A* protein, one of them (clone F1) was cloned into the polylinker region of the pTZ19 plasmid vector[14] to allow *in vitro*

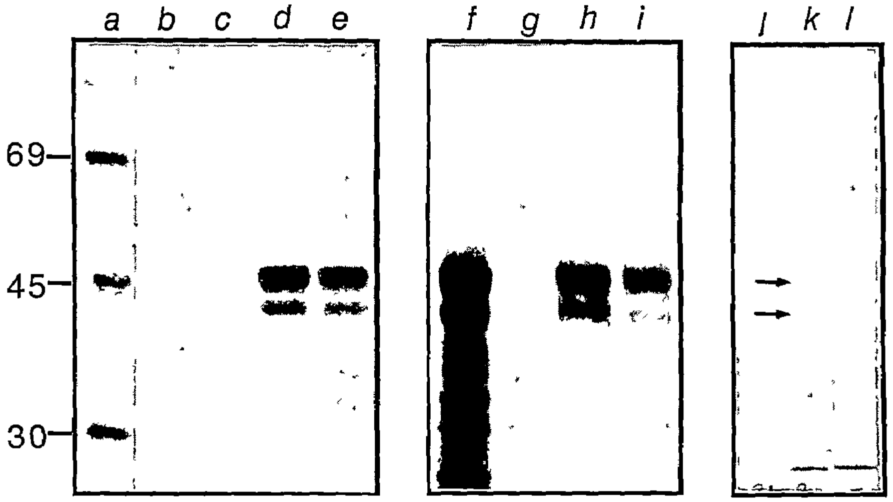


Fig. 2 Synthesis and immunoprecipitation of the c-*erb-A* protein. Capped mRNA, synthesized *in vitro* using T7 RNA polymerase, was translated in rabbit reticulocyte lysates. Lane *a*, marker proteins (sizes shown being relative molecular masses in thousands), lane *b*, unprogrammed lysate, lane *c*, translation with RNA from clone F1 containing the 288-nucleotides long 5′ untranslated sequence, lanes *d* and *e*, RNA from clones F1 and U1 with only 38 nucleotides of 5′ untranslated sequence remaining. Aliquots of the translation products were immunoprecipitated with two different anti-erb-A antisera. Total translation products with 5′-truncated RNA (lane *f*) were immunoprecipitated with preimmune serum (lane g) and with two different anti-erb-A antisera (lanes *h* and *i*). Chicken embryo fibroblasts, transiently transfected with expression plasmids, were labelled with [^{35}S]-methionine for 2 h. Immunoprecipitations using preimmune or anti-erb-A antiserum with lysates from cells receiving a c-*erb-A* expression plasmid are shown in lanes *i* and *k*, respectively. Immunoprecipitation with anti-erb-A antiserum of lysates from cells transfected with the plasmid lacking insert is shown in lane *l*.

Methods To reduce the length of the untranslated leader sequence, plasmids containing the entire F1 and U1 cDNA clones in the appropriate orientation in the *Eco*RI site of pTZ 19 were cleaved with *Eag*I 38 nucleotides upstream of the translation initiation codon, and with *Sma*I in the polylinker region of the vector, treated with the Klenow fragment of DNA polymerase I in the presence of all four deoxynucleotide triphosphates, and religated at low DNA concentration. For *in vitro* transcription, plasmids were linearized with *Hin*dIII or *Eco*RI, treated with proteinase K, extracted with phenol-chloroform and precipitated. Transcription was done with T7 DNA polymerase (100 units ml^{-1}, Biolabs) at 37 °C in 40 mM Tris *p*H 7.9, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM dithiothreitol (DTT), 0.5 mM of the four ribonucleotide triphosphates, 0.5 mM ^{7}mGpppG (Boehringer), 0.1 mg ml^{-1} bovine serum albumin (BSA), 1,000 U ml^{-1} Rnasin, at a DNA concentration of 0.1 mg ml^{-1}. Transcripts were translated in micrococcal-nuclease-treated rabbit reticulocyte lysates (NEN) in the presence of 3 mCi ml^{-1} [^{35}S]-methionine under conditions suggested by the manufacturer. Lysates were diluted 100-fold with lysis buffer before immunoprecipitation, kept on ice for 30 min and cleared by centrifugation at 15,000 r.p.m. for 30 min. Precipitations were essentially as described[3,51] using 5 μl of 1:10 diluted antiserum. Proteins were separated by SDS-polyacylamide gel electrophoresis on 10% gels, and the gels fluorographed and exposed at −70 °C. The anti-erb-A antisera were directed against the 5′ and 3′ domains of v-*erb-A* expressed in bacteria, and will be described in detail in a separate paper. To express c-*erb-A in vivo*, the 5′-truncated F1 clone was cloned into the expression plasmid pSV2 (ref. 15).

transcription with the T7 RNA polymerase and subsequent translation in reticulocyte lysates. In addition, to test if the 5′ untranslated sequence influenced the efficiency of translation, two constructs were made with cDNA from the clones F1 and U1, each lacking 250 nucleotides of this sequence. Figure 2 shows that RNA lacking the 5′ sequence (lanes *d* and *e*) directed the synthesis of a major 46K protein, whereas translation with RNA still containing this sequence gave only minor incorporation (lane *c*). Furthermore, the 46K protein as well as a minor 40K polypeptide were identified as c-*erb-A* proteins by immunoprecipitation with monospecific anti-v-erb-A antisera recognizing N- or C-terminal epitopes, respectively (Fig. 2, lanes *h* and *i*).

To demonstrate that the cDNAs directed the synthesis of c-*erb-A* protein also *in vivo*, chicken embryo fibroblasts were transiently transfected with a plasmid expressing c-*erb-A* from the simian virus 40 (SV40) promoter[15]. The cells were labelled after 48 h with [^{35}S]-methionine, lysed and subjected to immunoprecipitation with anti-erb-A antiserum. Figure 2 (lane *k*) shows that the c-*erb-A*-transfected cells contained 46 and 40K proteins, whereas no protein was specifically precipitated from cells transfected with the vector lacking an insert (lane *l*). The anti-erb-A antiserum also reacts with a 46K and a 40K protein expressed at very low levels in many types of avian cells (J.G., in preparation), suggesting that the polypeptides expressed from the cDNA represent authentic c-*erb-A* proteins.

The c-*erb-A* protein binds thyroid hormone

To identify a ligand for the c-*erb-A* protein we first searched the literature for a nuclear hormone receptor expressed in many types of cells that had M_r ~46K. The criteria were fulfilled by the nuclear receptor for the thyroid hormone triiodothyronine (T_3), as a 47K polypeptide that binds T_3 is expressed in many tissues[8,16,17]. Accordingly, c-*erb-A* protein produced by *in vitro* translation in reticulocyte lysates was incubated with [^{125}I]-labelled T_3 in the presence or absence of a 1,000-fold excess of unlabelled hormone, and tested for binding of hormone with a filter binding assay[18]. Figure 3*A* shows that lysates programmed with c-*erb-A* RNA containing the short 5′ untranslated sequence bound the radiolabelled hormone (column *a*, open), and that unlabelled T_3 (column *b*, open) or D-T_4, an analogue of T_3 which is biologically inactive but binds to the T_3 receptor[19], (column *c*, open) complete with most of this binding. In contrast, the binding exhibited by mock-programmed lysates in the presence or absence of competitor (stippled columns *a*, *b* and *c*) was similar to that determined for the c-*erb-A* lysate competed with unlabelled T_3, suggesting that the c-*erb-A* protein specifically binds T_3. Similar results (data not shown) were obtained when the binding was tested by chromatography on Sephadex G75 columns[20]. In addition, a 1,000-fold molar excess of unlabelled dexamethasone, oestradiol, progesterone, aldosterone, testosterone, vitamin D_3 (7-dehydrocholesterol), or tyrosine did not compete with radiolabelled T_3 for binding to c-*erb-A* containing lysates (open bars, Fig. 3*A*). Finally, an excess of these compounds did not compete with [^{125}I]-T_3 for nonspecific binding to mock-programmed lysates (compare the stippled columns *d*–*j* to *a*).

To demonstrate directly that the observed binding was due to c-*erb-A*, lysates containing *in vitro* translated c-*erb-A* protein were subjected to immunoadsorption with anti-erb-A or anti-erb-B antisera bound to protein A-Sepharose beads. The depleted samples were then assayed for binding of [^{125}I]-T_3 in the presence or absence of competitor. Column *m* in Fig. 3*A* shows that the adsorption with anti-erb-A antibodies removed about 70% of the specific binding activity, whereas extracts adsorbed with preimmune or anti-erb-B antisera exhibited no decreased

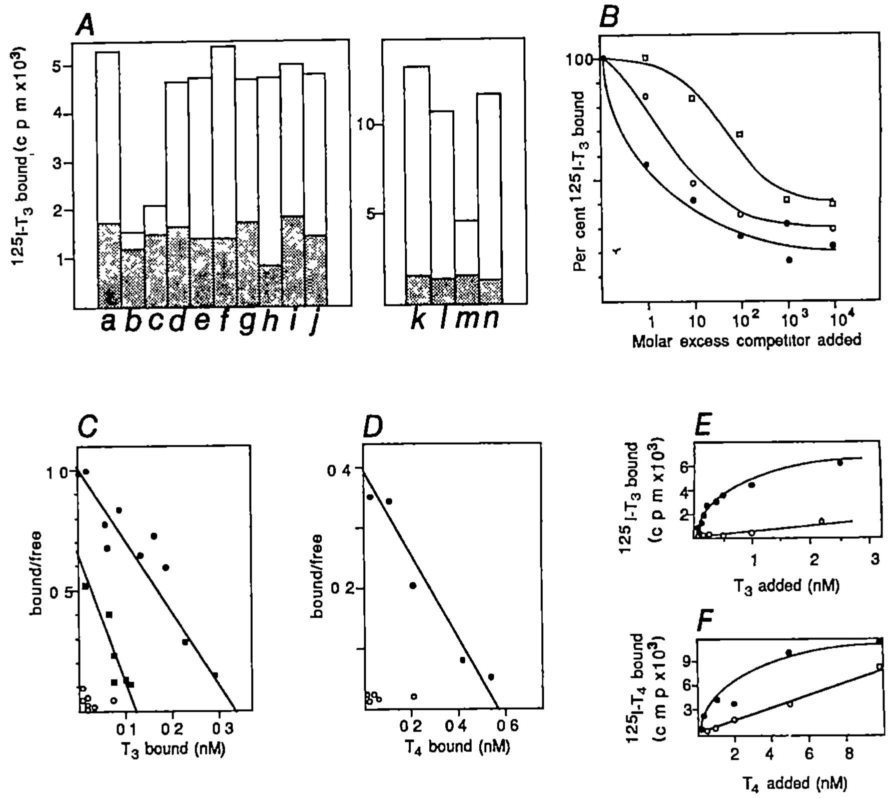


Fig 3 Ligand-binding properties of the c-*erb-A* protein produced by *in vitro* translation *A*, left, Reticulocyte lysates programmed with plus strand (open bars) or minus strand c-*erb-A* RNA (stippled bars) were incubated with 1 nM [^{125}I]-T$_3$ in the presence or absence of 1 μM of various unlabelled competitors *a*, no competitor, *b*, T$_3$, *c*, D-T$_4$, *d*, dexamethasone, *e*, oestradiol, *f*, progesterone, *g*, aldosterone, *h*, testosterone, *i*, vitamin D$_3$, *j*, tyrosine right, open bars show that the binding activity of c-*erb-A* protein for [^{125}I]-T$_3$ was reduced in lysates treated with anti-erb-A but not with anti-erb-B or preimmune sera Antisera *k*, no antibodies, *l*, pre-immune serum, *m*, anti-erb-A, *n*, anti-erb-B Stippled bars show binding in the presence of a 1,000-fold excess of unlabelled T$_3$ *B*, Competition binding with analogues of T$_3$ Lysates containing c-*erb-A* were incubated with 1 nM [^{125}I]-T$_3$ in the presence of the indicated molar excesses of unlabelled competitor Reverse T$_3$ (rT$_3$) is 3,3′,5′-thyronine ●, T$_3$, ○, T$_4$, □, rT$_3$ *C*, Scatchard analysis of T$_3$ binding Lysate extracts were diluted (●) 7-fold (K_d = 0 33 mM) or (■) 20-fold (K_d = 0 21 nM), ○, extract of mock-programmed lysate *D*, Scatchard analysis of T$_4$ binding ●, 7-fold diluted extract (K_d = 1 4 nM), ○, extract of mock-programmed lysate *e*, *f*, Saturation analysis of T$_3$ and T$_4$ binding Binding measured in extracts of (●) c-*erb-A* protein-containing or (○) mock-programmed lysates Nonspecific binding was determined by parallel incubations with 1 μM unlabelled T$_3$ or T$_4$

Methods All assays were done with protein synthesized *in vitro* essentially as described in Fig 2, except that DNA was purified on small Sepharose CL6B columns before transcription, no label was included in the translation reactions, and translation lysates were obtained from Amersham Aliquots of the lysate were incubated overnight at 0-4 °C with L-[3′-^{125}I]-triiodothyronine (3,000 mCi mg^{-1}) or L-[3′,5′-^{125}I]-thyroxine (1,500 mCi mg^{-1}, both Amersham) at the indicated concentrations (generally 0 5-2 0 μl lysate in a total volume of 10 μl) in binding buffer (50 mM NaCl, 2 mM EDTA, 5 mM mercaptoethanol, 20 mM Tris *p*H 7 5, 10% glycerol) Radiolabelled hormone was dried under vacuum to remove ethanol when the final isotope concentration used would exceed 2 nM Bound radiolabelled hormone was determined[18] by filtration at 4 °C using Millipore HAWP 02500 filters prewetted with 3 5 ml binding buffer Filters were washed by suction with 5 ml binding buffer at 4 °C three times, and were counted in a gamma counter Control lysates were programmed with RNA transcribed from a clone with the c-*erb-A* cDNA in the opposite orientation For immunodepletion, 5 μl of programmed lysate was diluted fivefold with binding buffer, incubated with 2 μl undiluted antiserum for 4 h at 4 °C, and adsorbed for 30 min to protein-A-Sepharose (saturated with erythroblast extract and washed extensively with binding buffer)

binding (columns *l* and *n*) Finally, the proteins adsorbed to the anti-erb-A antibodies bound no T$_3$, presumably due to conformational changes or steric hindrance imposed by the antibodies (data not shown)

The receptors for thyroid hormone bind with lower affinity analogues[21,22] of T$_3$, such as thyroxine (T$_4$) and reverse T$_3$ (rT$_3$) We therefore tested if higher concentrations of these analogues would be required to compete with labelled T$_3$ for binding to c-*erb-A* protein Figure 3*B* shows that T$_4$ was 5-10-fold and rT$_3$ was 50-100-fold less efficient than unlabelled T$_3$ in competing with [^{125}I]-T$_3$ These relative efficiencies are similar to those reported for competition of [^{125}I]-T$_3$ binding to the thyroid hormone receptors[21,22] and reflect the different biological activities of these products

To determine accurately the affinity of c-*erb-A* protein for thyroid hormone, we performed Scatchard analyses[23] with labelled T$_3$ and T$_4$ Because these hormones have been reported to bind with low affinity to unspecific sites at high protein concentrations[24] which affect the dissociation constants measured, we determined the K_d of c-erb-A for T$_3$ using reticulocyte lysates diluted to different concentrations Figure 3*C* (filled symbols) shows that the K_d for T$_3$ was between 0 21 and 0 33 nM with 20- and 7-fold diluted extracts, respectively, whereas the K_d for T$_4$, determined in a separate experiment with a 7-fold diluted extract, was 1 4 nM (Fig 3*D*) In contrast, Scatchard analyses done with mock-programmed lysates only indicated the presence of low-affinity/high-capacity binding sites (open symbols) The affinities of the c-*erb-A* protein for T$_3$ and T$_4$ are similar to the published values obtained with the thyroid hormone receptor in whole cells and in crude nuclear extracts[22,25]

Finally, the results from the saturation binding experiments (Fig 3*E* and *F*) showed that the high-affinity sites (closed symbols) were saturable with low concentrations of hormone, whereas the low-affinity sites present in the mock-programmed lysates were non-saturable (open symbols) By calculating the amount of [^{35}S]-methionine-labelled c-*erb-A* protein in the lysates and comparing to the amount of T$_3$ bound, we estimate that 5-10% of the total *in vitro* synthesized protein bound radiolabelled hormone, a result similar to those found for *in vitro* synthesized steroid receptors[26]

Cellular but not viral *erb-A* protein binds T$_3$

The T$_3$ receptor is nuclear regardless of whether or not it has bound hormone[8] To verify that the c-*erb-A* protein binds T$_3$ also *in vivo*, and to test the hormone binding capacity of p75$^{gag\text{-}v\text{-}erb\text{-}A}$, chicken embryo fibroblasts containing c-*erb-A* or v-*erb-A* were tested for hormone binding The c-*erb-A* gene was expressed from a replication-competent retrovirus vector[27], whereas a *wt* and an inactive v-*erb-A* gene (v-*erb-A2*$^-$, ref 1) were expressed from retrovirus constructs also conferring resistance to the drug G418 To measure nuclear T$_3$ binding *in vivo*, transfected and fully infected cell cultures were incubated with [^{125}I]-T$_3$ for three hours, and the radioactivity in the isolated nuclei determined Figure 4 (left) shows that the nuclei (open

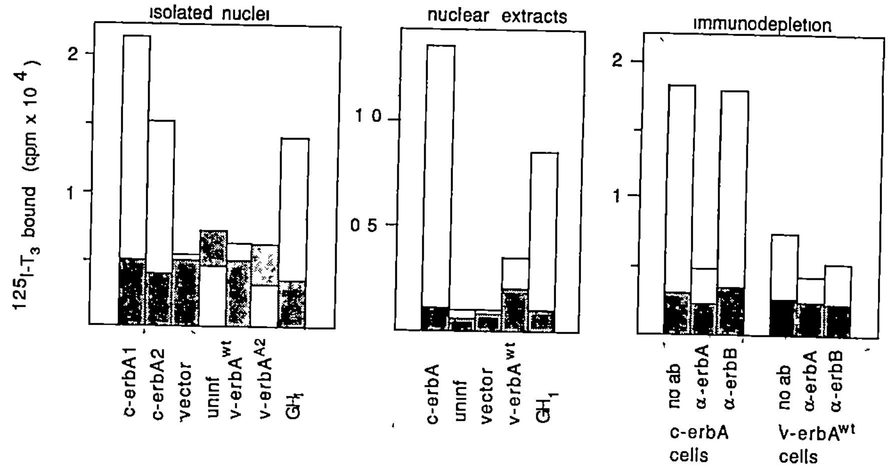


Fig 4 Hormone-binding properties of c-*erb-A* and v-*erb-A* proteins produced *in vivo* Left panel, Cells expressing c- or v-*erb-A* were incubated for 3 h with [^{125}I]-T$_3$, and the radioactivity in the nuclei determined Middle, Hormone binding by nuclear extracts incubated with [^{125}I]-T$_3$ *in vitro* Right panel, Reduction of *in vitro* binding activity in nuclear extracts by adsorption with anti-erb-A antisera Open bars show binding of radiolabelled hormone in the absence of competitor, whereas stippled bars indicate binding of [^{125}I]-T$_3$ in the presence of a 1,000-fold excess of unlabelled hormone

Methods The F1 cDNA clone containing only the short 5′ leader sequence was cloned into a modification of the replication competent avian retrovirus vector 779/989, a derivative of a previously described vector[27] Virus constructs expressing *wt* or mutant (v-*erb-A2*$^-$, ref 1) v-*erb-A* genes were constructed by excising the v-*erb-B* gene from AEV with *Apa*I and *Eco*RI, followed by appropriate modification and insertion of the *Bgl*III–*Bam*HI fragment of the Tn*5 neo*^{R} element The constructs were introduced into chicken embryo fibroblast cells by the $Ca_3(PO_4)_2$ precipitation technique[52 53] In the case of the c-*erb-A* construct, virus infection was monitored by reverse transcriptase activity in the culture media, whereas cells containing the v-*erb-A-neo*^{R} viruses were superinfected with RAV1 helper virus and then isolated after selection for G418 resistance Control experiments showed that these viruses expressed v-*erb-A* at levels similar to those found in AEV-transformed cells (not shown) Cells were grown in Eagle's medium plus 5% fetal calf serum and 2% heat-inactivated chicken serum For analysis of nuclear T$_3$-receptor content, cells grown for 24–48 h in media containing T$_3$-depleted[54] serum were trypsinized, washed twice with Hank's and resuspended at ~3×10^6 cells ml^{-1} in T$_3$-free medium supplemented with 0 15 nM [^{125}I]-T$_3$ for 3 h Cells were subsequently washed with Hank's, and lysed in 1 ml 50 mM Tris *p*H 7 85, 1 1 mM $MgCl_2$, 0 5% Triton-X-100 at 0 °C Nuclei were collected by centrifugation (1,500 r p m for 5 min) and were washed once more in lysis buffer before counting GH$_1$ cells were cultured in Eagle's medium containing 10% horse serum and 2 5% fetal calf serum Isolation of nuclei and preparation of nuclear extracts were done employing techniques used for isolation of nuclear T$_3$ receptor[55], except that inhibitors of proteolysis were included in all steps of isolation Experiments on [^{125}I]-T$_3$ binding with nuclear extracts and immunodepletion were done as described in Fig 3

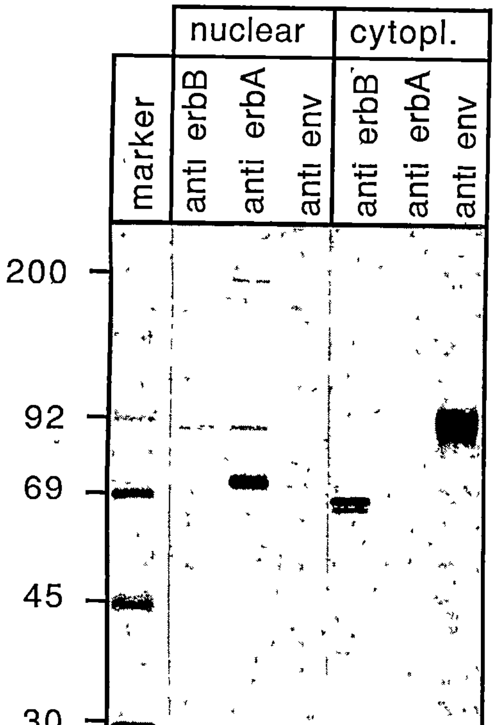


Fig 5 Demonstration of a nuclear localization for the v-*erb-A* protein Cell fractionation experiments with AEV-transformed erythroblasts were done as described in Fig 4 Labelling of cells and immunoprecipitations were carried out as described in Fig 2 Sizes for markers are relative molecular masses in thousands

columns) from two different cultures of c-*erb-A* expressing cells (c-*erb-A*1 and 2) bound [^{125}I]-T$_3$ at levels similar to that found in the nuclei of GH$_1$ cells, a rat pituitary line expressing 1–2×10^4 T$_3$ receptors per cell[28,29] In addition, a 1,000-fold excess of unlabelled T$_3$ competed with this binding (stippled columns) In contrast, only a low amount of specific T$_3$ binding was found in the nuclei from the *wt* and the mutant v-*erb-A* expressing cells, or the cells infected with the replication competent vector lacking c-*erb-A* insert Similar results were obtained when nuclear extracts, prepared from unlabelled cells, were tested for binding of [^{125}I]-T$_3$ *in vitro* (middle panel) Again, receptors from the c-*erb-A* and the control GH$_1$ cells bound radioactivity at levels 8–12-fold higher (open bars) than that observed when binding was done in the presence of an excess of unlabelled T$_3$ (stippled bars) In contrast, the binding seen with extracts from uninfected cells or cells expressing v-*erb-A* was only 1 3–1 7 times background Lastly, adsorption with anti-erb-A antisera of nuclear extracts from c-*erb-A* but not v-*erb-A* cells reduced the specific binding by 77% (Fig 4, right panel), demonstrating that the high T$_3$ binding was dependent on c-*erb-A* protein

To exclude the possibility that lack of nuclear hormone binding by v-*erb-A*-expressing cells was due to cytoplasmic localization of p75$^{gag\text{-}v\text{-}erb\text{-}A}$, a cell fractionation experiment was done Erythroblasts transformed by AEV were labelled with [^{35}S]-methionine, and nuclear and cytoplasmic extracts, prepared using techniques described for the isolation of the T$_3$ receptor[17], were immunoprecipitated with antisera against v-erb-A, v-erb-B and the viral envelope glycoprotein Figure 5 shows that the v-*erb-A* protein was present predominantly in the nuclear extract, whereas the v-*erb-B* and *env* proteins partitioned in the cytoplasmic fraction Previous reports have suggested that v-*erb-A* is localized in the cytosol[30], these cell fractionation experiments were however done under conditions that may not preserve the nuclear localization of at least the unoccupied oestradiol receptor[31,32]

Discussion

Several lines of evidence suggest that the avian c-*erb-A* gene codes for a thyroid hormone receptor the molecular weight of the c-*erb-A* protein is the same as that of a previously characterized receptor, it binds T$_3$ and its analogues with affinities typical of the T$_3$ receptor, binding activity can be removed with anti-erb-A antibodies, and both the thyroid hormone receptor and c-erb-A are expressed at low levels in many tissues Finally, Weinberger and collaborators have demonstrated that a human c-*erb-A* protein also binds thyroid hormone *in vitro*[33] However, the evidence is so far based solely on hormone-binding experiments, and definite proof will require the demonstration that the c-*erb-A* protein exerts the same effects *in vivo* as the thyroid hormone receptor

Thyroid hormone induces a wide variety of metabolic effects, for instance increased lipogenesis and ketogenesis These effects have been studied in great detail in liver cells but have also been observed in many other types of tissues[8] The hormone also plays key roles in differentiation, such as the rapid development of frogs from thyrectomized tadpoles, the induction of growth hormone production in the pituitary, and the development of the brain in neonatals by promoting dendrite formation and myelination[8] Interestingly, T$_3$ does not induce the same physiological effects in the brain as in other tissues[34], suggesting that the brain receptor could be distinct from the receptor(s) found in other cells In fact, several c-*erb-A* related genes appear to exist in the human genome A c-*erb-A* gene highly homologous to the avian gene reported in this paper is located on human chromosome 17 (refs 35-38), whereas other distinct genes have been found on chromosome 3 (ref 33)

The thyroid hormone receptor also has a nuclear localization in the unoccupied state as well, and is thought to act by binding to DNA (refs 39, 40) thus inducing or suppressing gene expression in a manner similar to that of steroid receptors[41] The v-*erb-A* protein binds no hormone, yet it is nuclear, binds to DNA *in vitro* (H B and U Gehring, unpublished data) and exerts specific physiological effects in erythroblasts[3] It is therefore likely that p75$^{gag\text{-}v\text{-}erb\text{-}A}$ has lost its hormone-binding capacity as a result of one or several of the mutations it has accumulated Moreover, it is unclear whether it affects the expression of the same genes and in the same manner as the normal receptor, as its presumed DNA-binding region has two amino-acid substitutions It is possible that p75$^{gag\text{-}v\text{-}erb\text{-}A}$, although unable to bind ligand, acts in a similar fashion as the increased nuclear transfer (nt^{i}) mutants previously described for the glucocorticoid receptor[41]

The maturation of erythroid cells is accompanied by synthesis of several components important for proper red cell differentiation, such as band 3, the major anion transporter which also serves as an anchor for erythrocyte cytoskeletal proteins, and hemin, a possible regulatory factor of differentiation[42,43] Interestingly, thyroid hormone regulates the induction of key liver enzymes involved in hemin synthesis[44], and induces increased levels of the Na$^+$/K$^+$ ATPase[45,46] We are therefore currently investigating the possibility that v-erb-A exerts its effects in erythroblasts by suppressing or inducing the synthesis of ion transporters and hemin

It has previously been demonstrated that thyroid status has a modulating effect on neoplasia Administration of thyroid hormone to thyrectomized rodents is a prerequisite for the induction of hepatomas by certain chemicals, indicating a role in the initiating action of the carcinogen In contrast, thyroid hormone appears to promote the growth and metastatic potential of other types of neoplasias, demonstrating an enhancement of tumour progression[47] In addition, the transformation of cultured cells by radiation or adenovirus infection is facilitated by thyroid hormone[48] The v-*erb-A* oncogene of AEV enhances the transforming activity of other oncogenes in erythroblasts but also in fibroblasts (refs 1-3, and M Jansson and B V, unpublished data) In addition, the integration of a hepatitis B viral genome close to a human *erb-A* related gene was recently reported[49] The demonstration that the c-*erb-A* protein has properties similar to those of the T$_3$ receptor may thus represent the first identification of an oncogene that enhances oncogenesis without being able to induce detectable neoplastic effects on its own

We thank Dr Kristina Nordstrom for valuable help and Dr Thomas Graf for helpful suggestions We thank Dr Steve Hughes for providing the replication competent retrovirus vector, and Dr Kåre Gautvik for the gift of GH$_1$ cells A M was supported by a fellowship from the Juan March Foundation during sabbatical leave from Consejo Superior de Investigationes Cientificas, Spain

Received 23 October, accepted 13 November 1986

1 Frykberg, L *et al* *Cell* **32**, 227-238 (1983)
2 Kahn, P *et al* *Cell* **45**, 349-356 (1986)
3 Damm, K, Beug, H, Graf, T & Vennstrom, B *EMBO J* (submitted)
4 Weinberger, C, Hollenberg, S M, Rosenfeld, M G & Evans, R M *Nature* **318**, 670-672 (1985)
5 Green, S *et al* *Nature* **320**, 134-139 (1986)
6 Krust, A *et al* *EMBO J* **5**, 891-897 (1986)
7 Conneely, O M *et al* *Science* **233**, 767-770 (1986)
8 Oppenheimer, J H & Samuels, H H (eds) *Molecular Basis of Thyroid Hormone Action* (Academic, New York, 1983)
9 Vennstrom, B & Bishop, J M *Cell* **28**, 135-143 (1982)
10 Watson, C J & Jackson, J F in *DNA cloning* (ed Glover, D M) Vol 1, 79-88 (IRL, Oxford, 1985)
11 Young, R A & Davis, R W *Proc natn Acad Sci USA* **80**, 1194-1198 (1983)
12 Debuire, B *et al* *Science* **224**, 1456-1459 (1984)
13 Schwartz, D E., Tizard, R & Gilbert, W *Cell* **32**, 853-869 (1983)
14 *Analects* 13, 1-2 (Pharmacia Uppsala, 1986)
15 Mulligan, R C & Berg, P *Science* **209**, 1422-1427 (1980)
16 Pascual, A, Casanova, J & Samuels, H H *J biol Chem* **257**, 9640-9647 (1985)
17 Dozin, B, Cahnmann, H J & Nikodem, V M *Biochemistry* **24**, 5197-5202 (1982)
18 Inoue, A, Yamakawa, J, Yakioka, M & Morisawa, S *Analyt Biochem* **134**, 176-183 (1983)
19 Samuels, H H, Stanley, F & Casanova, J *J clin Invest* **63**, 1229-1240 (1979)
20 Samuels, H H, Tsai, H S & Casanova, J *Science* **184**, 1188-1191 (1974)
21 Oppenheimer, J H in *Molecular Basis of Thyroid Hormone Action* (eds Oppenheimer, J H & Samuels, H H) 1-34 (Academic, New York 1983)
22 Latham, K R, Ring, J C & Baxter, J D *J biol Chem* **251**, 7388-7397 (1976)
23 Scatchard, G *Ann N Y Acad Sci* **51**, 660-672 (1949)
24 Seelig S, Schwartz H L & Oppenheimer, J H *J biol Chem* **256**, 2154-2161 (1981)
25 Samuels, H H in *Molecular Biology of Thyroid Hormone Action* (eds Oppenheimer, J H & Samuels, H) 36-65 (Academic, New York, 1983)
26 Kumar, V, Green, S, Staus, A & Chambom, P *EMBO J* **5**, 2231-2236 (1986)
27 Hughes, S & Kosik, E *Virology* **136**, 89-99 (1984)
28 Samuels, H H, Stanley, F & Shapiro, L E *Proc natn Acad Sci USA* **73**, 3877-3881 (1976)
29 Samuels, H H, Stanley, F & Shapiro, L E *J biol Chem* **252**, 6052-6060 (1977)
30 Bunte, T, Greiser-Wilke, I, Donner, P & Molling, K *EMBO J* **1**, 919-927 (1982)
31 King, W J & Greene, G L *Nature* **302**, 745-747 (1984)
32 Welshons, W V Liebermann M E & Gorski, J *Nature* **307**, 747-749 (1984)
33 Weinberger, C *et al* *Nature* **324**, 641-646 (1986)
34 Wilson, J D & Foster, D W (eds) *Williams Textbook of Endocrinology* (Saunders, Philadelphia, 1985)
35 Jansson M, Philipson, L & Vennstrom, B *EMBO J* **2**, 561-565 (1983)
36 Spurr, N *et al* *EMBO J* **3**, 159-163 (1984)
37 Dayton, A *et al* *Proc natn Acad Sci USA* **81**, 4495-4499 (1984)
38 Zabel, B U *et al* *Proc natn Acad Sci USA* **81**, 4874-4878 (1984)
39 Macleod, K M & Baxter, J D *J biol Chem* **251**, 7380-7387 (1976)
40 Wilson, B D, Wium, C A & Gent, W L *Biochem biophys Res Commun* **124**, 29-36 (1984)
41 Yamamoto, K A *A Rev Genet* **19**, 209-252 (1985)
42 Fukuda, M, Fukuda, M N, Papayannopoulou, T & Hakomori, S-I *Proc natn Acad Sci USA* **77**, 3474-3478 (1980)
43 Mager, D & Bernstein, A *J cell Physiol* **100**, 467-480 (1979)
44 Sassa, S, Bradlow, H L & Kappas, A *J biol Chem* **254**, 10011-10020 (1979)
45 Lo, C S & Edelman, I S *J biol Chem* **251**, 7834-7840 (1976)
46 Lo, C-S & Lo, T N *J biol Chem* **255**, 2132-2136 (1980)
47 Guernsey, D L, Borek, C, Fisher, P B & Edelman, I S in *Genes and Proteins in Oncogenesis* (eds Weinstein, I B & Vogel, H J), (Academic, New York, 1983)
48 Fisher, P B, Guernsey, D L, Weinstein, I B & Edelman, I S *Proc natn Acad Sci USA* **80**, 196-200 (1983)
49 Dejean, A, Bougueleret, L, Grzeschik, K-H & Tiollais, P *Nature* **322**, 70-72 (1986)
50 Sanger, F, Nicklen, S & Coulson, A R *Proc natn Acad Sci USA* **74**, 475-477 (1977)
51 Hayman, M *et al* *Cell* **32**, 579-588 (1983)
52 Graham, F L & van der Eb, A J *Virology* **52**, 456-467 (1973)
53 Stow, N D & Wilkie, N W *J gen Virol* **33**, 447-453 (1976)
54 Samuels, H H, Stanley, F & Casanova, J *Endocrinology* **105**, 80-85 (1978)
55 Samuels, H H & Tsai, J S *Proc natn Acad Sci USA* **70**, 3488-3492 (1973)

The c-*erb-A* gene encodes a thyroid hormone receptor

Cary Weinberger*, Catherine C. Thompson*†, Estelita S. Ong*, Roger Lebo‡, Donald J. Gruol§ & Ronald M. Evans*

* Howard Hughes Medical Institute, Gene Expression Laboratory, and § Regulatory Biology Laboratory, The Salk Institute for Biological Studies, PO Box 85800, San Diego, California 92138, USA
† Department of Biology, University of California, San Diego, California 92093, USA
‡ Howard Hughes Medical Institute, Department of Medicine, University of California, San Francisco, California 94143, USA

*The cDNA sequence of human c-*erb-A*, the cellular counterpart of the viral oncogene v-*erb-A*, indicates that the protein encoded by the gene is related to the steroid hormone receptors Binding studies with the protein show it to be a receptor for thyroid hormones*

THE interaction of DNA-binding proteins with specific DNA sequences is one mechanism by which gene transcription is controlled Characterization of the regulatory proteins, which include steroid hormone receptors, should help our understanding of DNA–protein interactions and of the control of gene activation and repression The human glucocorticoid receptor (hGR) complementary DNA has been sequenced[1,2] and shown to be functionally active[3] Interestingly, sequence analysis of the receptor showed it to be related to the product of the v-*erb-A* oncogene product of avian erythroblastosis virus (AEV)[4] This led to the proposal that the steroid receptors and the *erb-A* oncogene products share a common primordial archetype and that the *erb-A* proto-oncogene products may also be proteins that bind to DNA enhancer elements Recent characterization of the human oestrogen[5,6], chicken progesterone[7,8] and human aldosterone (J Arriza, C W and R M E, unpublished data) receptors further support these conclusions

Accordingly, we started characterizing the human c-*erb-A* proto-oncogene even though its functional identification could not be ensured During the progress of these studies, advances were made in detailing the functional domains of the glucocorticoid receptor[3] which indicated that the hGR hormone binding domain was unusually large, encompassing the carboxy-terminal 300 amino acids This entire region has distant but significant similarity to the carboxy terminus of v-erb-A which therefore focused our attention on classes of molecules that might exert transcriptional regulatory effects similar to those of steroid hormones

The molecular mechanism of thyroid hormone stimulation of gene expression may be similar to that outlined for steroids[9] Thyroid hormone is present in all chordate species examined and exerts profound effects on development and differentiation, such as metamorphosis in amphibians[9] Like steroids, thyroid hormones may enter cells by passive diffusion and bind to high affinity nuclear receptors which in turn mediate a rapid and selective activation of gene expression[10,11] Evidence favouring this hypothesis has come largely from studies of the induction of growth hormone and its messenger RNA in the somatotroph of the rat anterior pituitary and in a number of related rat somatotrophic cell lines[12,13] Thyroid hormones rapidly increase the transcription of the growth hormone gene in these cells[14 15] The increase in transcription is accompanied by increased levels of nuclear thyroid hormone–receptor complexes, is time- and concentration-dependent, and is independent of protein synthesis[16-18]

The similarity of steroid and thyroid hormone actions led us to examine the possibility that the *erb-A* protein may itself be the thyroid hormone receptor We first isolated and characterized a cDNA for human c-*erb-A* The sequence predicts a 456 amino-acid polypeptide containing a cysteine/lysine/arginine-rich region similar to the putative DNA-binding domain of steroid hormone receptors and a carboxy-terminal region distantly related to the steroid binding domain Using a functional assay developed for analysing the hormone binding properties of cloned steroid receptors[2], we demonstrated that the translation product from the human c-*erb-A* cDNA possesses intrinsic thyroid hormone binding activity, characteristic of the native thyroid hormone receptor molecule

Characterization of c-*erb-A* cDNAs

To isolate a human c-*erb-A* cDNA, a 500-base pair (bp) *Pst*I DNA fragment isolated from a region of the AEV genome containing only the v-*erb-A* gene[19] was used as a ^{32}P-labelled probe to screen two human placenta cDNA libraries Two overlapping λgt10 cDNA clones were obtained by screening two independent libraries of ~10^6 phage recombinants each The restriction maps were deduced for each of the cDNA clone *Eco*RI inserts from pUC8 subclones, pheA4 and pheA12 (Fig 1*a*) Nucleotide sequence analysis of these overlapping 1 5 kilobase (kb) cDNA clones revealed that the composite sequence contains a long open reading frame of 456 amino acids with a presumptive initiator methionine codon at nucleotide number 301 and a terminator codon at 1,669 in the sequence (Fig 1*b*) Seven codons upstream of the ATG is an in-frame terminator TAA providing support for the initiator methionine, although another methionine found 26 codons downstream makes this assignment tentative No consensus polyadenylation addition signal (AATAAA, ref 20) is discernible in the 27 nucleotides between the terminator and poly(A) tract in pheA12

A predicted polypeptide of relative molecular mass 52,000 ($M_r = 52$K), encoded within the human c-*erb-A* cDNA translational open reading frame, shares 82% amino acid identity with the region downstream from the viral *gag* sequences[21] in AEV No gaps in the amino acid comparison were revealed The human c-*erb-A* amino acid sequence is homologous with the viral protein beginning at viral amino acid residue 37 (Fig 2) The carboxy-terminus of c-*erb-A* differs from that of v-*erb-A* amino acid sequence similarity of the two polypeptides terminates at residue 445 of c-erb-A and residue 380 of v-erb-A (Fig 2)

Alignment of the human c-*erb-A* nucleic acid sequence with that of v-*erb-A* shows that ~74% of the human gene is identical to the viral gene in the region of amino acid homology (nucleotides 563–1636 in c-*erb-A*, data not shown) These comparisons, coupled with previous data describing other human *erb-A* genes mapping to chromosome 17 (refs 22–24), indicate that the human placenta c-*erb-A* gene we have isolated is distinct One of the two chromosome 17 *erb-A* genes (both of which we propose to call hc-*erb-A*α) has 82% similarity with the v-*erb-A* gene by nucleotide sequence and 89% by amino acid identity[24], There-

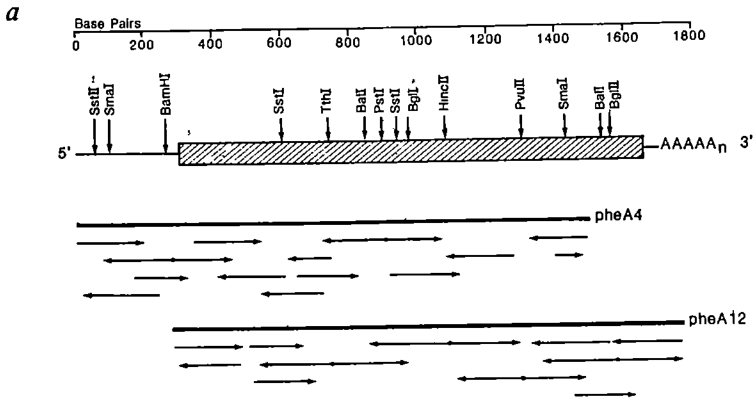


Fig. 1 Restriction map and sequencing strategy (*a*) and nucleotide and predicted amino acid sequence (*b*) of human placenta c-*erb-A* cDNA *a*, Orientations of the two subclones pheA4 and pheA12 relative to the composite restriction map Common restriction endonuclease cleavage sites are above the linear map Thin lines, untranslated sequences, hatched box, *erb-A* coding region, arrows, DNA fragments sequenced *b*, Nucleotide sequence of the composite *erb-A* cDNA is presented in the 5' to 3' orientation The translational open reading frame related to the viral erb-A protein[21] is shown above the nucleotide sequence Adenosine residues (~130) are found at the 3' end of pheA12 Numbers above the translated sequence indicate amino acid residues and nucleotide numbers are on the right of the sequence **Methods** Recombinant phage (~10^6) from each of two human placenta λgt10 cDNA libraries[44] were screened using a nick-translated[45] 500-bp *Pst*I fragment isolated from pAEV-11 (ref 19) The hybridization mixture contained 50% formamide, 1× Denhardt's, 5× SSPE, 0 1% sodium dodecyl sulphate (SDS), 100 $\mu g\ ml^{-1}$ denatured salmon sperm DNA and 10^6 c p m ml^{-1} of ^{32}P-labelled *Pst*I fragment (specific activity = 1×10^8 c p m μg^{-1}) Duplicate nitrocellulose filters were hybridized at 37 °C for 18 h, washed three times for 20 min each in 0 1× SSC, 0 1% SDS (1×SSC = 150 mM NaCl, 15 mM trisodium citrate) and autoradiographed at −70 °C with an intensifying screen Two hybridization-positive clones were isolated, subcloned into the *Eco*RI site of pUC8, and sequenced by the chemical cleavage method[46]

b

```
CGGCGGGGATCAACTTTGCATGAATAATGTGAGTGCGCTTGGAAAAGAGACCTCCTGCTCCGCGGGCTCGGGGCAAGAGCCCGCAGGCTA   90

CCTTCCCCGGGCAGGGGCGCTCAACCCAACCGGCTCCAGGGCACTGGTAATTTGGCTAGAGGACCGCGCGGAGGCAGCGGGATCTGCGAT  180

TTCCTTCTGGTTGGCTGTCCTGCGTGGGTGCCAAGTTCCACACATGATTTAATGAATAAGAAGGAGATGTCAGTGAAAAAAGGGATCCAG  270

                                                         10                            20
                              MetThrGluAsnGlyLeuThrAlaTrpAspLysProLysHisCysProAspArgGluHis
AATGATTACTAACCTATAACCCCCAACAGTATGACAGAAAATGGCCTTACAGCTTGGGACAAACCGAAGCACTGTCCAGACCGAGAACAC  360

                           30                            40                            50
AspTrpLysLeuValGlyMetSerGluAlaCysLeuHisArgLysSerHisSerGluArgArgSerThrLeuLysAsnGluGlnSerSer
GACTGGAAGCTAGTAGGAATGTCTGAAGCCTGCCTACATAGGAAGAGCCATTCAGAGAGGCGCAGCACGTTGAAAAATGAACAGTCGTCG  450

                           60                            70                            80
ProHisLeuIleGlnThrThrTrpThrSerSerIlePheHisLeuAspHisAspAspValAsnAspGlnSerValSerSerAlaGlnThr
CCACATCTCATCCAGACCACTTGGACTAGCTCAATATTCCATCTGGACCATGATGATGTGAACGACCAGAGTGTCTCAAGTGCCCAGACC  540

                           90                           100                           110
PheGlnThrGluGluLysLysCysLysGlyTyrIleProSerTyrLeuAspLysAspGluLeuCysValValCysGlyAspLysAlaThr
TTCCAAACGGAGGAGAAGAAATGTAAAGGGTACATCCCCAGTTACTTAGACAAGGACGAGCTCTGTGTAGTGTGTGGTGACAAAGCCACC  630

                           120                           130                           140
GlyTyrHisTyrArgCysIleThrCysGluGlyCysLysGlyPhePheArgArgThrIleGlnLysAsnLeuHisProSerTyrSerCys
GGGTATCACTACCGCTGTATCACGTGTGAAGGCTGCAAGGGTTTCTTTAGAAGAACCATTCAGAAAAATCTCCATCCATCCTATTCCTGT  720

                           150                           160                           170
LysTyrGluGlyLysCysValIleAspLysValThrArgAsnGlnCysGlnGluCysArgPheLysLysCysIleTyrValGlyMetAla
AAATATGAAGGAAAATGTGTCATAGACAAAGTCACGCGAAATCAGTGCCAGGAATGTCGCTTTAAGAAATGCATCTATGTTGGCATGGCA  810

                           180                           190                           200
ThrAspLeuValLeuAspAspSerLysArgLeuAlaLysArgLysLeuIleGluGluAsnArgGluLysArgArgArgGluGluLeuGln
ACAGATTTGGTGCTGGATGACAGCAAGAGGCTGGCCAAGAGGAAGCTGATAGAGGAGAACCGGGAGAAAAGACGGCGGGAAGAGCTGCAG  900

                           210                           220                           230
LysSerIleGlyHisLysProGluProThrAspGluGluTrpGluLeuIleLysThrValThrGluAlaHisValAlaThrAsnAlaGln
AAGTCCATCGGGCACAAGCCAGAGCCCACAGACGAGGAATGGGAGCTCATCAAAACTGTCACCGAAGCCCATGTGGCGACCAACGCCCAA  990

                           240                           250                           260
GlySerHisTrpLysGlnLysProLysPheLeuProGluAspIleGlyGlnAlaProIleValAsnAlaProGluGlyGlyLysValAsp
GGCAGCCACTGGAAGCAAAAACCGAAATTTCTGCCAGAAGACATTGGACAAGCACCAATAGTCAATGCCCCAGAAGGTGGAAAGGTTGAC 1080

                           270                           280                           290
LeuGluAlaPheSerHisPheThrLysIleIleThrProAlaIleThrArgValValAspPheAlaLysLysLeuProMetPheCysGlu
TTGGAAGCCTTCAGCCATTTTACAAAAATCATCACACCAGCAATTACCAGAGTGGTGGATTTTGCCAAAAAGTTGCCTATGTTTTGTGAG 1170

                           300                           310                           320
LeuProCysGluAspGlnIleIleLeuLeuLysGlyCysCysMetGluIleMetSerLeuArgAlaAlaValArgTyrAspProGluSer
CTGCCATGTGAAGACCAGATCATCCTCCTCAAAGGCTGCTGCATGGAGATCATGTCCCTTCGCGCTGCTGTGCGCTATGACCCGGAAAGT 1260

                           330                           340                           350
GluThrLeuThrLeuAsnGlyGluMetAlaValIleArgGlyGlnLeuLysAsnGlyGlyLeuGlyValValSerAspAlaIlePheAsp
GAGACTTTAACCTTGAATGGGGAAATGGCAGTGATACGGGGCCAGCTGAAAAATGGGGGTCTTGGGGTGGTGTCAGACGCCATCTTTGAC 1350

                           360                           370                           380
LeuGlyMetSerLeuSerSerPheAsnLeuAspAspThrGluValAlaLeuLeuGlnAlaValLeuLeuMetSerSerAspArgProGly
CTAGGCATGTCTCTGTCTTCTTTCAACCTGGATGACACTGAAGTAGCCCTCCTTCAGGCCGTCCTGCTGATGTCTTCAGATCGCCCGGGG 1440

                           390                           400                           410
LeuAlaCysValGluArgIleGluLysTyrGlnAspSerPheLeuLeuAlaPheGluHisTyrIleAsnTyrArgLysHisHisValThr
CTTGCCTGTGTTGAGAGAATAGAAAAGTACCAAGATAGTTTCCTGCTGGCCTTTGAACACTATATCAATTACCGAAAACACCACGTGACA 1530

                           420                           430                           440
HisPheTrpProLysLeuLeuMetLysValThrAspLeuArgMetIleGlyAlaCysHisAlaSerArgPheLeuHisMetLysValGlu
CACTTTTGGCCAAAACTCCTGATGAAGGTGACAGATCTGCGGATGATAGGAGCCTGCCATGCCAGCCGCTTCCTGCACATGAAGGTGGAA 1620

                           450
CysProThrGluLeuLeuProProLeuPheLeuGluValPheGluAspEnd
TGCCCCACAGAACTCCTCCCCCCTTTGTTCCTGGAAGTGTTCGAGGATTAGACTGACTGGATTCCTTCCTATAATTCCAAAAAAAAAAAA 1710
```

fore, the placenta c-*erb-A* cDNA, which we propose to call hc-*erbAβ*, is more distantly related to the viral *erb-A* gene than the hc-*erbAα* genes

Amino acid sequence comparisons between the viral and cellular *erb-A* protein products and the glucocorticoid receptor indicate graded levels of homology with the carboxy-terminal half of the hGR (Fig 2) The highest degree of similarity is found in a cysteine-rich sequence of 65 amino acids beginning at c-*erb-A* amino acid residue 102 (ref 4) There is 47% amino acid identity in the comparison with the hGR and 52% identity when c-*erb-A* is compared with the human oestrogen receptor (hER) amino acid sequence (Fig 2) We have proposed that this region of the hGR represents the DNA binding domain[4] Mutagenesis and expression studies have provided direct evidence for its role in transcriptional activation[3] Regions downstream from the cysteine-rich domain, which correspond to the hormone binding domain, contain reduced yet significant (17%) homology to the hGR and hER as found previously with the viral *erb-A* product[4-6]

Multiple *erb-A* genes

Hybridization of restriction endonuclease-digested human placenta DNA with a labelled DNA fragment derived from the cysteine-rich region of the c-*erb-A* polypeptide (Fig 1*a*,*b*) pro-

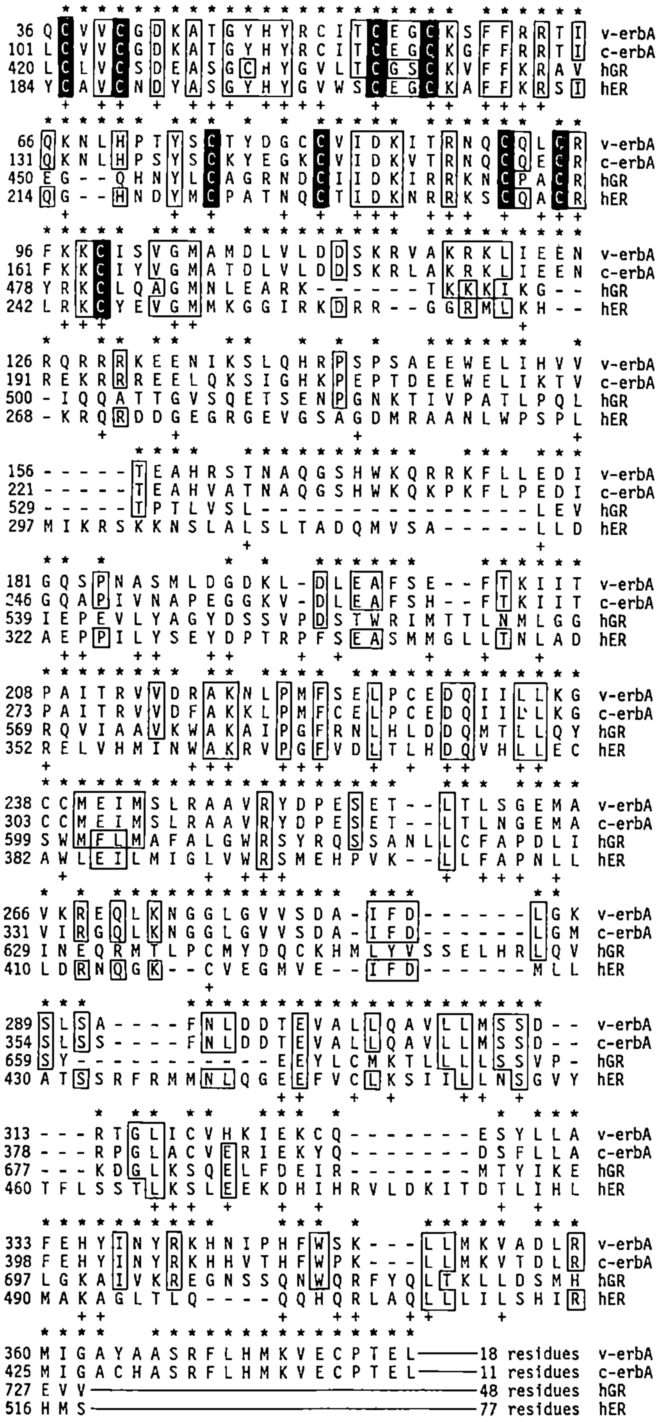

36 Q C V V C G D K A T G Y H Y R C I T C E G C K S F F R R T I v-erbA
101 L C V V C G D K A T G Y H Y R C I T C E G C K G F F R R T I c-erbA
420 L C L V C S D E A S G C H Y G V L T C G S C K V F F K R A V hGR
184 Y C A V C N D Y A S G Y H Y G V W S C E G C K A F F K R S I hER
66 Q K N L H P T Y S C T Y D G C C V I D K I T R N Q C Q L C R v-erbA
131 Q K N L H P S Y S C K Y E G K C V I D K V T R N Q C Q E C R c-erbA
450 E G - - Q H N Y L C A G R N D C I I D K I R R K N C P A C R hGR
214 Q G - - H N D Y M C P A T N Q C T I D K N R R K S C Q A C R hER
96 F K K C I S V G M A M D L V L D D S K R V A K R K L I E E N v-erbA
161 F K K C I Y V G M A T D L V L D D S K R L A K R K L I E E N c-erbA
478 Y R K C L Q A G M N L E A R K - - - - - - T K K K I K G - - hGR
242 L R K C Y E V G M M K G G I R K D R R - - G G R M L K H - - hER
126 R Q R R R K E E N I K S L Q H R P S P S A E E W E L I H V V v-erbA
191 R E K R R R E E L Q K S I G H K P E P T D E E W E L I K T V c-erbA
500 - I Q Q A T T G V S Q E T S E N P G N K T I V P A T L P Q L hGR
268 - K R Q R D D G E G R G E V G S A G D M R A A N L W P S P L hER
156 - - - - - T E A H R S T N A Q G S H W K Q R R K F L L E D I v-erbA
221 - - - - - T E A H V A T N A Q G S H W K Q K P K F L P E D I c-erbA
529 - - - - - T P T L V S L - - - - - - - - - - - - - - - L E V hGR
297 M I K R S K K N S L A L S L T A D Q M V S A - - - - - L L D hER
181 G Q S P N A S M L D G D K L - D L E A F S E - - F T K I I T v-erbA
246 G Q A P I V N A P E G G K V - D L E A F S H - - F T K I I T c-erbA
539 I E P E V L Y A G Y D S S V P D S T W R I M T T L N M L G G hGR
322 A E P P I L Y S E Y D P T R P F S E A S M M G L L T N L A D hER
208 P A I T R V V D R A K N L P M F S E L P C E D Q I I L L K G v-erbA
273 P A I T R V V D F A K K L P M F C E L P C E D Q I I L L K G c-erbA
569 R Q V I A A V K W A K A I P G F R N L H L D D Q M T L L Q Y hGR
352 R E L V H M I N W A K R V P G F V D L T L H D Q V H L L E C hER
238 C C M E I M S L R A A V R Y D P E S E T - - L T L S G E M A v-erbA
303 C C M E I M S L R A A V R Y D P E S E T - - L T L N G E M A c-erbA
599 S W M F L M A F A L G W R S Y R Q S S A N L L C F A P D L I hGR
382 A W L E I L M I G L V W R S M E H P V K - - L L F A P N L L hER
266 V K R E Q L K N G G L G V V S D A - I F D - - - - - - - L G K v-erbA
331 V I R G Q L K N G G L G V V S D A - I F D - - - - - - - L G M c-erbA
629 I N E Q R M T L P C M Y D Q C K H M L Y V S S E L H R L Q V hGR
410 L D R N Q G K - - C V E G M V E - - I F D - - - - - - - M L L hER
289 S L S A - - - - F N L D D T E V A L L Q A V L L M S S D - - v-erbA
354 S L S S - - - - F N L D D T E V A L L Q A V L L M S S D - - c-erbA
659 S Y - - - - - - - - - - - - E E Y L C M K T L L L L S S V P - hGR
430 A T S S R F R M M N L Q G E E F V C L K S I I L L N S G V Y hER
313 - - - R T G L I C V H K I E K C Q - - - - - - - - E S Y L L A v-erbA
378 - - - R P G L A C V E R I E K Y Q - - - - - - - - D S F L L A c-erbA
677 - - - K D G L K S Q E L F D E I R - - - - - - - - M T Y I K E hGR
460 T F L S S T L K S L E E K D H I H R V L D K I T D T L I H L hER
333 F E H Y I N Y R K H N I P H F W S K - - - L L M K V A D L R v-erbA
398 F E H Y I N Y R K H H V T H F W P K - - - L L M K V T D L R c-erbA
697 L G K A I V K R E G N S S Q N W Q R F Y Q L T K L L D S M H hGR
490 M A K A G L T L Q - - - - Q Q H Q R L A Q L L L I L S H I R hER
360 M I G A Y A A S R F L H M K V E C P T E L 18 residues v-erbA
425 M I G A C H A S R F L H M K V E C P T E L 11 residues c-erbA
727 E V V 48 residues hGR
516 H M S 77 residues hER

Fig 2 Amino acid sequence comparison between the carboxy-terminal portions of the v-*erb-A* oncogene product, the human placenta c-*erb-A* polypeptide and the human glucocorticoid and oestrogen receptors Translated amino acid sequences for both the v-*erb-A* protein (upper sequence) and the human placenta c-*erb-A* polypeptide (second sequence) were compared by aligning matching residues A computer program for the concurrent comparison of three or more amino acid sequences[47] was used to align c-*erb-A*, human glucocorticoid receptor (hGR third sequence from top)[2] and human oestrogen receptor (hER, bottom sequence)[5] carboxy-terminal amino acid sequences, on the basis of progressive evaluation of selected segments from each sequence Amino acid residues matched in at least three of the polypeptides are boxed Amino acid matches between the two erb-A polypeptides are indicated by an asterisk above the top sequence in each column Amino acid identities between the steroid receptors are designated by crosses below the sequences Hyphens, gaps inserted to maximize the no of matches in the comparison Cysteine residues conserved between the four polypeptides printed white-on-black

duced two bands in every digestion with the exception of *Pvu*II (Fig 3*a*). The greater intensity of the 9 4 kb-band suggested that it contains two hybridizing DNA fragments When the hybridization conditions were relaxed, additional bands were observed in the products of each enzyme digestion For example, two faint bands of 5 1 and 3 6 kb were seen after *Pst*I digestion (Fig 3*b*) The hybridization probe contains a single internal *Pst*I site (Fig 1*a*) which probably explains the increased number of *Pst*I bands detected with this probe

Similar high-stringency hybridization experiments were performed using a 260 bp *Eco*RI-*Bam*HI fragment from the 5' untranslated region of pheA4 (Fig 1*a* and data not shown) Two hybridizing DNA fragments were detected with all restriction enzymes providing further support for the existence of two related c-*erb-A* genes No additional DNA bands were seen when the hybridization was performed under relaxed conditions using this probe (data not shown) We conclude that there are two closely related hc-*erb-Aβ* proto-oncogenes and a third, less similar one, in the human genome

Laser-sorted chromosomes were prepared from human lymphoid cells[25], bound to nitrocellulose filters and hybridized under non-stringent hybridization and washing conditions using the 1 5-kb *Eco*RI insert from pheA4 (Fig 3*c*) This probe detected only human chromosome 3-specific DNA and suggests that the three c-*erb-Aβ*-related genes are chromosomally linked, although we cannot strictly exclude the possibility that the non-stringently hybridizing *erb-A* gene maps in another chromosome Interestingly, another steroid receptor/*erb-A* genomic fragment has recently been identified by characterizing the integration site for hepatitis B virus in a human hepatocellular carcinoma[26] This locus has also been mapped on human chromosome 3 suggesting that the *erb-A* genes may be closely linked

Expression of c-*erb-A* genes

Northern blots hybridizing cytoplasmic poly(A)-containing RNAs, isolated from various human cell lines or human term placenta, with a 650-bp *Bam*HI-*Pst*I fragment from pheA4 (Fig 1*a*) revealed a single RNA species of 2,000 nucleotides that is most abundant in HeLa and MCF-7 cells (Fig 4*a*) The size of the mRNA indicates that we have isolated a nearly full-length c-*erb-A* cDNA HT1080 cells contain a small amount of the 2-kb transcript while it is undetectable in IM-9 cells Human placenta appears to contain multiple species of 5, 3 and 2 5 as well as 2 0 kb RNA It is unclear whether the multiple placenta bands represent nuclear precursors, or mature mRNAs from a single gene, or the products of other *erb-A* genes

The protein products of the human c-*erb-A* cDNA were characterized by *in vitro* translation A cDNA containing the entire c-*erb-A* coding region was inserted into the *Eco*RI site of the expression vector pGEM3 in both orientations Capped RNA transcripts synthesized by T7 polymerase from these templates were used to programme protein synthesis in rabbit reticulocyte lysates and the ^{35}S-methionine-labelled products were separated on SDS-polyacrylamide gels[27] Proteins with M_r of 55, 52, and 35K were detected when peA101 DNA (*erb-A* sense transcripts) was used as template (Fig 4*b*, lanes 3 and 4) but not when peA102 DNA (*erb-A* antisense transcripts) was used (Fig 4*b*, lane 2) The 55 and 52K products may correspond to polypeptides initiating translation at methionines 1 and 27, respectively (Fig 1), while the 35K product may be a proteolytic breakdown product

Thyroid hormone binding

Structural similarity of the steroid ligands as well as the partial amino acid sequence homology (40%) between the carboxy termini of the hGR and hER (which specify the hormone binding domains, ref 28) support the hypothesis that the steroid receptors comprise a family of regulatory proteins The more distant

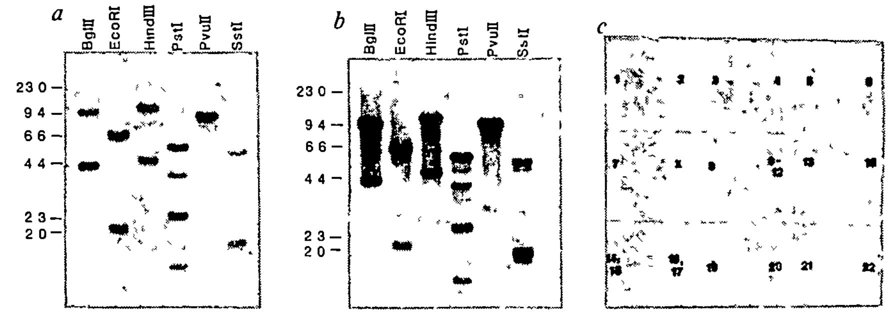


Fig. 3 Southern analysis and chromosome mapping of human placenta DNA with c-*erb-A* DNA probes. *a*, Human term placenta DNA was digested with restriction endonucleases and products were separated on a 0.8% agarose gel. DNAs were transferred to nitrocellulose paper[48] and hybridized in 50% formamide, 5× SSPE, 1× Denhardt's, 0.1% SDS, 100 μg ml^{-1} salmon sperm DNA with the 450-bp *Sst*I fragment from pheA4 which was nick-translated to a specific activity of 5×10^{8} c.p.m. μg^{-1}. The filter was washed in 0.1× SSC, 0.1% SDS at 60 °C and exposed to X-ray film at −70 °C with an intensifying screen. Lambda *Hin*dIII DNA markers (size in kb) are aligned to the left of the autoradiogram. *b*, Analysis of placenta DNA using the same c-*erb-A* probe as in *a* under non-stringent hybridization conditions. A parallel blot containing the same samples was hybridized as in *a*, except that 35% formamide was used. The hybridized filter was washed in 2× SSC, 0.1% SDS at 55 °C and exposed to X-ray film as described above. *c*, Chromosome mapping of human c-*erb-A* genes. Human lymphocyte chromosomes were separated by laser cytofluorometry[25] and probed using the non-stringent hybridization conditions described in *b*, but with the 1.5 kbp *Eco*RI insert from pheA4 as probe.

homology between the carboxy terminus of c-*erb-A* and the steroid receptors suggested that *erb-A* proto-oncogenes probably do not encode steroid receptors, but is consistent with the hypothesis that *erb-A* may respond to other molecules. We have shown that *in vitro* translation can be used as a means to characterize hormone binding activity of cloned steroid receptors[2]. Accordingly, this assay was employed to identify the putative c-*erb-A* ligand.

Steroid and thyroid hormones exert their effects through fundamentally similar mechanisms. To address the possibility that erb-A may be the thyroid hormone receptor, the *in vitro* translation products were mixed with ^{125}I-3,5,3′-triiodo-L-thyronine (^{125}I-T_3) in a hormone binding reaction[29]. Non-specific hormone binding was determined by adding a 500-fold molar excess of cold T_3 to parallel samples. Remarkably, the mixture containing the 55 and 52K polypeptides (peA101) acquired ^{125}I-T_3 binding whereas the anti-sense RNA-programmed lysates (peA102) had only background binding. Hormone binding was sensitive to proteases but not nucleases (data not shown). Affinity of T_3 for the cloned *erb-A* protein was determined by Scatchard analysis (Fig. 5*a*). A K_d value of 5×10^{-11} M was obtained, which is virtually identical to the T_3 binding (6×10^{-11} M) in HeLa cell nuclear extracts (data not shown).

Specific analogues of thyroid hormones have characteristic competition patterns for T_3 binding to the native thyroid hormone receptor. We determined whether the *erb-A* product synthesized *in vitro* shared the same intrinsic hierarchy of affinities for a range of natural and synthetic thyroid hormones. Competition of ^{125}I-T_3 binding was most effectively achieved with 3,5′, 3′-triiodothyroacetic acid (TRIAC) which inhibited 50% of ^{125}I-T_3 binding at 300 pM (Fig. 5*b*). In addition, L-thyroxine (Fig. 5*b*), D-T_3 and reverse T_3 (3,3′,5′-triiodo-L-thyronine) competed more poorly than T_3 (Fig. 5*c*), while 100 μM vitamin D_3, aldosterone, cortisol, testosterone, progesterone or oestradiol did not compete (data not shown), consistent with the biochemical properties of the rat thyroid hormone receptor[29–31].

High salt (0.4 M KCl) HeLa cell nuclear extracts contained thyroid hormone (^{125}I-T_3) binding activity, while none was found in cytoplasmic or low-salt (0.1 M KCl) nuclear extracts (data not shown). Competition of thyroid hormone binding in the nuclear extracts was quantitatively similar to that of lysates containing c-erb-A made *in vitro* (compare Fig. 5*b* and *d*). Furthermore, thyroid hormone binding using 0.4 M KCl nuclear extracts from IM-9 cells, which contain undetectable levels of c-*erb-A* mRNA (Fig. 4, lane 4), is negligible when compared with similar HeLa extracts (data not shown). These results provide direct evidence that c-erb-A is the thyroid hormone receptor.

Conclusions

The data in this report provide three criteria that identify the c-*erb-A* product as the thyroid hormone receptor. First, the overall structural homology of c-erb-A to steroid hormone receptors suggests that erb-A is likely to be a ligand-responsive regulatory protein. Second, the expressed protein product has the same intrinsic hierarchy of affinities for natural and synthetic thyroid hormones as the native receptor. Third, the molecular weights of erb-A *in vitro* translation products are similar to the photo-affinity-labelled rat thyroid hormone receptor[39]. The identity of erb-A as the thyroid hormone receptor could be further sub-

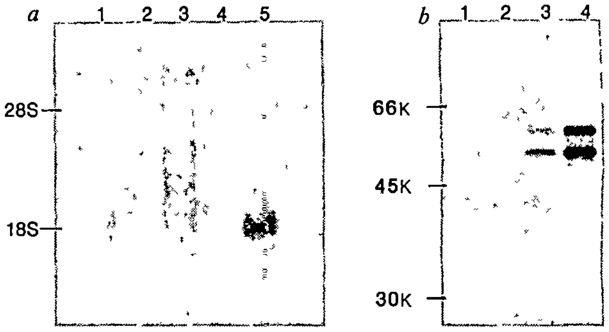


Fig. 4 Human c-*erb-A* expression. *a*, Northern analysis of RNAs from human cell lines and human placenta. Cytoplasmic poly(A)-containing RNA (12 μg) from HeLa, MCF-7 and IM-9 cells, or total poly(A) RNAs from HT1080 or placenta, were separated on a 1% agarose gel containing formaldehyde, transferred to nitrocellulose[49] and probed using a nick-translated 650-bp *Bam*HI-*Pst*I pheA4 fragment. Cytoplasmic RNAs were isolated from the cell lines using isotonic buffer and 0.5% NP40, while the placenta RNAs were extracted from fresh tissue using guanidine thiocyanate[50]. Lane 1, HeLa; lane 2, HT1080; lane 3, human placenta; lane 4, IM-9; lane 5, MCF-7. *b*, Synthesis of erb-A polypeptides *in vitro*. The ^{35}S-methionine-labelled products synthesized using T7 polymerase-catalysed RNA transcripts were separated on a 12.5% SDS-polyacrylamide gel which was fluorographed (EN^{3}HANCE, NEN). Lane 1, control (without mRNA); lane 2, peA102 (anti-sense RNA, 4 μl); lane 3, peA101 (sense RNA, 1 μl); lane 4, peA101, (4 μl RNA). Sizes of protein standards: bovine serum albumin, 66.2K; ovalbumin, 45K and carbonic anhydrase, 31K.

Methods. The *Eco*RI insert from pheA12 which contains the entire coding region of c-*erb-A* was inserted into the *Eco*RI site of pGEM3 (Promega Biotec) in both orientations. Plasmid DNAs peA101 and peA102 were linearized with *Hin*dIII, purified on 0.8% agarose gels and used as templates for T7 polymerase-catalysed synthesis of RNAs *in vitro*[2]. After P60 chromatography, total nucleic acid material (2 μg) was used to programme protein translation in a rabbit reticulocyte lysate system (Promega Biotec) with ^{35}S-methionine (25 μCi, 1100 Ci mmol^{-1}, NEN) final volume 25 μl.

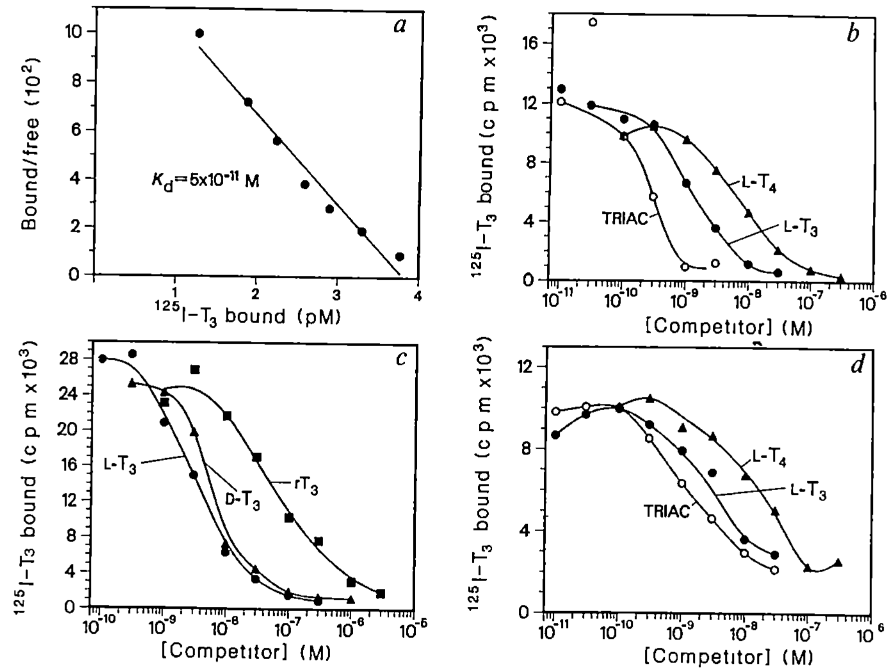


Fig 5 Thyroid hormone binding to erb-A polypeptides synthesized *in vitro* *a*, Scatchard analysis of ^{125}I-T$_3$ binding to the erb-A polypeptides made *in vitro* The erb-A polypeptides (2 μl from the *in vitro* translation mixture in a total volume of 2 ml) were assayed for specific thyroid hormone binding activity using hydroxylapatite to measure the amount of bound and free labelled hormone at different concentrations of ^{125}I-T$_3$ as described[51] *b*, Competition of thyroid hormone analogues for ^{125}I-T$_3$ binding to erb-A polypeptides synthesized *in vitro* Samples (2 μl) from the peA101(sense strand)-programmed reactions were used in ^{125}I-T$_3$ standard binding reactions[29] with increasing concentrations of unlabelled thyroid hormone or analogues to compete with labelled hormone, Specifically bound thyroid hormone is plotted against concentration of competitor compound *c*, Competition of triiodothyronine isomers for ^{125}I-T$_3$ binding to erb-A polypeptides synthesized *in vitro* Binding reactions were performed as above adding increasing concentrations of T$_3$ isomers Thyroid hormone bound to erb-A is plotted on the ordinate *d*, Competition of thyroid hormone analogues for ^{125}I-T$_3$ binding to 0 4 M KCl HeLa cell nuclear extracts HeLa cell nuclei were extracted with a buffer containing 0 4 M KCl (ref 29) The protein extract (25 μg) (determined by Bio-Rad protein assay) was mixed with 0 6 nM ^{125}I-T$_3$ in standard binding reactions with increasing concentrations of thyroid hormone and analogues[29 30]

Methods Labelled ^{125}I-3,3′, 5-triiodo-L-thyronine (NEN, 2,200 Ci mmol^{-1}, 0 3 nM final) was mixed with erb-A polypeptides synthesized in the *in vitro* translation mixture (described in Fig 4) in T$_3$-binding buffer (0 25 M sucrose, 0 25 M KCl, 20 mM Tris-HCl (*p*H 7 5), 1 mM MgCl$_2$, 2 mM EDTA, 5 mM dithiothreitol (DTT))[29] at 0 °C for 2 h in a final volume of 250 μl Specific hormone binding was determined by adding a 1,000-fold excess of unlabelled hormone and assayed by counting radioactivity eluting in the excluded volume from a Sephadex G-25 fine (Pharmacia) 0 9×4 0 cm column[29]

stantiated by demonstrating its transcriptional regulation of T$_3$-responsive genes such as the growth hormone gene

Analysis of the hGR and hER has revealed the proteins to be composed of a series of functional domains[2-4,32-35] These include a cysteine-rich region which contains structural similarity to a repeated cysteine-rich sequence found in *Xenopus* 5S gene transcription factor IIIA and other transcriptional regulatory proteins[36,37] as well as carboxy-terminal sequences, which encode the steroid-binding domain[3,4,38] Extension of this analogy to the thyroid hormone receptor would predict its hormone binding region to be localized near the carboxy-terminal end of the molecule (Fig 6) Putative DNA binding sequences would be found in the cysteine-rich region (Fig 6) where DNA binding properties of the hGR and hER appear to be localized (refs 3, 38, S Hollenberg and R M E , unpublished data)

Thyroid hormone receptor and oncogenesis

Expression of the v-*erb-A* product in avian erythroblasts is required for maintenance of the fully transformed phenotype[40-43] Chickens infected with viruses lacking the v-*erb-A* gene display a less virulent disease, while *in vitro* these infected erythroblasts differentiate spontaneously and grow only with complex media supplements Cells infected with *erb-A*$^+$/*erb-B*$^+$ virus, however, have an increased capacity for self-renewal and display a less differentiated phenotype Structural alterations of the v-*erb-A* protein could give rise to a product exerting aberrant properties of growth control For instance, changes at the carboxyl terminus might affect thyroid hormone binding activity, as has been shown for the *β* form of the human glucocorticoid receptor, resulting in a constitutively active molecule[2-4], where changes abolish steroid binding activity Insertional mutants in this domain also inactivate steroid binding properties[3] However, deletion of the hormone binding region gives rise to a constitutively active receptor indicating that this domain plays a modulatory role in transcriptional activation (V Giguere and R M E , unpublished data) These data lead us to predict that v-erb-A is unlikely to bind hormone and is rather a constitutively active form of the thyroid hormone receptor The identification of erb-A as the thyroid hormone receptor provides the first direct evidence of a causative involvement of enhancers and their binding proteins in oncogenic transformation

A superfamily of regulatory genes

Similarity of the steroid receptors with the v-*erb-A* oncogene product was sufficient to allow us to propose that both have evolved from a primordial receptor gene[4] Two surprising results have emerged from the studies presented here The first is that

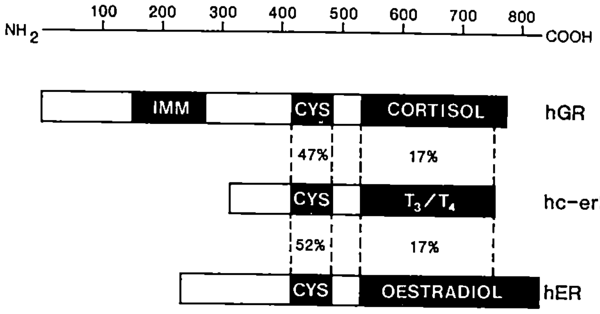


Fig 6 Schematic comparison of the steroid and thyroid hormone receptors Amino-acid sequences of the receptor molecules aligned in Fig 2 are represented schematically CYS, cysteine-rich region encoding the putative DNA binding domain found in the receptor proteins (Cys-rich region residues are c-erb-A, 102–169, hGR, 421–486, hER, 185–250) Cortisol, oestradiol and T$_3$/T$_4$), hormone binding regions in the carboxyl termini, IMM, immunogenic region of the human glucocorticoid receptor Numbers separating boxes, percentage amino acid identities between the receptor species in the intervals between the vertical broken lines, hGR, human glucocorticoid receptor, hER human oestrogen receptor, hc-erb-A*β*, human thyroid hormone receptor

the occurrence of a family of *erb-A* proto-oncogenes implies the existence of one or more other molecules closely related to the thyroid hormone receptor Physiological studies have not predicted the existence of a second class of thyroid hormone receptors and thus the characterization of this family may shed new light on mechanisms of developmental and homeostatic regulation The second surprising observation from these results is the close kinship of the thyroid hormone receptors with the steroid hormone receptor family This relationship indicates that these molecules may all be part of a superfamily of regulatory proteins that have arisen over evolutionary time to match the increasing developmental and physiological demands of more complex eukaryotes Thus, the putative identification of the *erb-A* protein as the thyroid hormone receptor leads to new insight into these issues and should provide an important tool for analysis of the molecular basis of transcriptional activation by regulatory proteins

We thank Dr Michael G Rosenfeld for discussions, Gail Cerelli and Kevin Murakami for technical help during the initial phase, Drs Bart Sefton and Richard Heyman for suggestions, and Drs Russell F Doolittle and Mark Johnson for carboxy-terminal sequence alignments C W thanks Robin for helpful discussions We also thank Marijke ter Horst for artwork and help in the preparation of the manuscript C W and R L are research associates of the Howard Hughes Medical Institute C T is a pre-doctoral trainee, supported by a training grant to the Department of Biology, University of California, San Diego R M E is an investigator of the Howard Hughes Medical Institute This work was supported by grants from the NIH and the Howard Hughes Medical Institute

Received 20 October, accepted 19 November 1986

1 Weinberger, C *et al.* *Science* **228**, 740-742 (1985)
2 Hollenberg, S M *et al.* *Nature* **318**, 635-641 (1985)
3 Giguere, V, Hollenberg, S M, Rosenfeld, M G & Evans, R M *Cell* **46**, 645-642 (1986)
4 Weinberger, C, Hollenberg, S M, Rosenfeld, M G & Evans, R M *Nature* **318**, 670-672 (1985)
5 Green, S *et al.* *Nature* **320**, 134-139 (1986)
6 Greene, G L *et al* *Science* **231**, 1150-1154 (1986)
7 Jeltsch, J M *Proc. natn. Acad. Sci. U S A.* **83**, 5424-5428 (1986)
8 Conneely, O M *et al.* *Science* **233**, 767-770 (1986)
9 Eberhardt, N L, Apriletti, J W & Baxter, J D in *Biochemical Actions of Hormones* Vol VII (ed Litwack, G) 311-394 (Academic, New York, 1980)
10 Tata, J R. & Widnell, C C *Biochem J* **98**, 604-620 (1966)
11 Oppenheimer, J H, Koerner, D, Schwartz, H L & Surks, M I J *Clin. Endocr Metab* **35**, 330-333 (1972)
12 Samuels, H H, Tsai, J S & Cintron, R. *Science* **181**, 1253-1256 (1973)
13 Tsai, J S & Samuels, H H *Biochem biophys Res Commun* **59**, 420-428 (1974)
14 Martial, J A, Baxter, J D, Goodman, H M & Seeburg, P H *Proc. natn. Acad. Sci U S A.* **74**, 1816-1820 (1977)
15 Evans, R. M, Birnberg, N C & Rosenfeld, M G *Proc. natn Acad. Sci. U S A.* **79**, 7659-7663 (1982)
16 Samuels, H H & Shapiro, L. E *Proc. natn. Acad Sci. U S A.* **73**, 3369-3373 (1976)
17 Spindler, B J, Mellon, S H & Baxter, J D *J biol. Chem.* **257**, 11627-11632 (1982)
18 Yaffe, B M & Samuels, H H *J biol Chem.* **259**, 6284-6291 (1984)
19 Vennstrom, B, Fanshier, L, Moscovici C & Bishop J M *J Virol.* **36**, 575-585 (1980)
20 Proudfoot, N J & Brownlee, G G *Nature* **263**, 211-214 (1976)
21 Debuire, B *et al.* *Science* **224**, 1456-1459 (1984)
22 Jansson, M, Philipson, L & Vennstrom, B *EMBO J* **2**, 561-565 (1983)
23 Spurr, N K *et al.* *EMBO J* **3**, 159-163 (1984)
24 Dayton, A I *Proc. natn. Acad Sci U S A.* **81**, 4495-4499 (1984)
25 Lebo, R. *et al.* *Science* **225**, 57-59 (1984)
26 Dejean, A, Bougueleret, L, Grzeschik, K-H & Tiollais, P *Nature* **322**, 70-72 (1986)
27 Laemmli, U K *Nature* **227**, 680-685 (1970)
28 Krust, A *et al.* *EMBO J* **5**, 891-897 (1986)
29 Samuels, H H, Tsai, J S, Casanova, J & Stanley, F *J clin Invest* **54**, 853-865 (1974)
30 Samuels, H H, Stanley, F & Casanova, J *J clin. Invest* **63**, 1229-1240 (1979)
31 Latham, K R, Ring, J C & Baxter, J D *J biol Chem.* **251**, 7388-7397 (1976)
32 Carlstedt-Duke, J, Okret, S, Wrange, O & Gustafsson, J-A *Proc. natn Acad Sci. U S A.* **79**, 4260-4264 (1982)
33 Dellweg, H G, Hotz, A, Mugele, K. & Gehring, U *EMBO J* **1**, 285-289 (1982)
34 Reichman, M E *et al.* *Biochemistry* **23**, 5376-5384 (1984)
35 Sherman, M R, Pickering, L A, Rollwagen, F M & Miller, L K. *Fedn Proc.* **37**, 167-73 (1978)
36 Miller, J, McLachlan, A D & Klug, A *EMBO J* **4**, 1609-1614 (1985)
37 Berg, J *Science* **232**, 485-487 (1985)
38 Kumar, V, Green, S, Staub, A & Chambon, P *EMBO J* **5**, 2231-2236 (1986)
39 Pascual, A, Casanova, J & Samuels, H H *J biol. Chem.* **257**, 9640-9647 (1982)
40 Graf, T & Beug, H *Cell* **34**, 7-9 (1983)
41 Frykberg, L *et al.* *Cell* **32**, 227-238 (1983)
42 Sealy, L, Privalsky, M L, Moscovici, G, Moscovici, C & Bishop, J M *Virology* **130**, 155-178 (1983)
43 Kahn, P *et al.* *Cell* **45**, 349-356 (1986)
44 Huynh, T V, Young, R A & Davis, R. W in *DNA Cloning, a Practical Approach* Vol 1 (ed Glover, D) 49-78 (IRL Press, Oxford, 1985)
45 Rigby, P W J, Dieckmann, M, Rhodes, C & Berg, P *J molec. Biol.* **113**, 237-251 (1977)
46 Maxam, A & Gilbert, W *Proc. natn. Acad. Sci. U S A.* **74**, 560-564 (1977)
47 Johnson, M S & Doolittle, R F *J molec. Evol.* **23**, 267-278 (1986)
48 Southern, E M *J molec. Biol.* **98**, 503-517 (1975)
49 Thomas, P *Proc. natn Acad Sci. U S A.* **77**, 5201-5205 (1980)
50 Chirgwin, J M, Przybyla, A F, MacDonald, R J & Rutter, W F *Biochemistry* **18**, 5294-5299 (1979)
51 Gruol, D J *Analyt Biochem* **101**, 387-393 (1980)

LETTERS TO NATURE

Rotational period of comet Halley

R. L. Millis & D. G. Schleicher

Lowell Observatory, 1400 West Mars Hill Road, Flagstaff,
Arizona 86001, USA

The rotational period of the nucleus of comet Halley is a basic datum which must be accurately known if the vast collection of observations of this comet recently obtained from spacecraft and ground-based observatories are to be correctly interpreted. Sekanina and Larson, who have studied the morphology of the comet's coma, find a period near 2.2 days, as do investigators associated with the Planet-A, Vega 1 and Vega 2 imaging experiments[1-4]. Others, however, have reported significantly different values, or have been unable to find clear evidence of periodicity in their data[5-7]. We report here the results of ground-based photometry of comet Halley carried out during March and April 1986 which indicate that the rotational period of the comet is, in fact, near 7.4 days.

During an extended photometric study of comet Halley during the 1985-86 apparition, two lengthy observing runs were scheduled on the 0 61-m Lowell-Tololo telescope at Cerro Tololo Inter-American Observatory (CTIO) in March and April of 1986 The March run (4-18 March) was timed to span the encounters of the Vega, Planet-A, and Giotto spacecraft with the comet, while the April run (29 March-19 April) covered the time of the comet's closest approach to Earth Of the 37 nights on which we had use of the telescope during the March and April runs, 36 were, at least in part, photometric Consequently, much high-quality photometric data were obtained While our main aim was to measure production rates of various gases and dust and to determine their variation with heliocentric distance, it quickly became apparent that the comet was undergoing sizeable night-to-night changes in activity

The observations were made using a conventional single-channel photometer equipped with pulse-counting electronics Interference filters, mostly chosen from the standard IAU set[8], were used to isolate emission bands of OH, NH, CN, C_3, CO^+, and C_2 along with continuum regions at 3,650 and 4,845 Å Observations were made with the comet's apparent photocentre centred in the entrance aperture of the photometer Each observation typically consisted of two or three 10-s integrations on the comet minus a sky measurement made several degrees away Extinction corrections were determined nightly and have been applied in the standard fashion The observed intensities were then transformed to absolute fluxes by reference to nightly observations of several standard stars Finally, the underlying

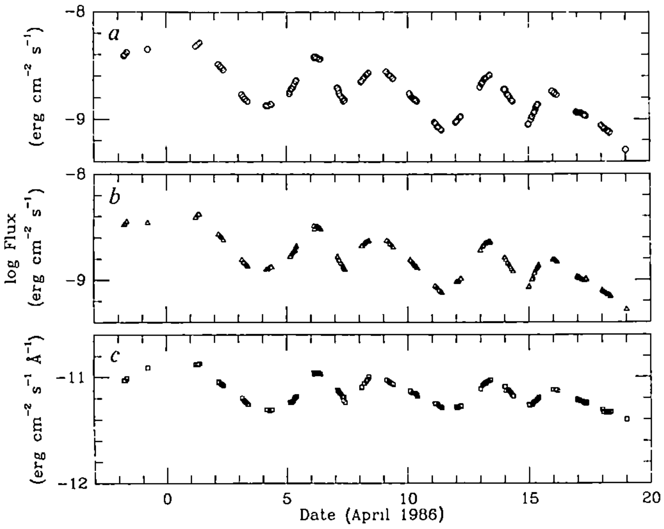


Fig. 1 Brightness variations in the inner coma of comet Halley seen during April 1986 in the light of emission bands of *a*, C_2 (5,140 Å), *b*, CN (3,870 Å), *c*, in a continuum passband at 3,650 Å.

continuum within the emission-band filters, based on the continuum-band results, was subtracted.

Figure 1 shows the flux from comet Halley through selected passbands measured during the April run. These observations were made with a photometer entrance aperture of 39-arc s radius. Light curves of essentially identical shape were recorded through the other filters and with entrance apertures of different sizes. The observational uncertainty in each plotted point is comparable to or less than the size of the point. It is evident from a brief inspection of Fig. 1 that the brightness of the central region of the coma was varying by a factor of ~4 on a timescale of a few days.

For brevity, we will discuss only observations made in a single passband centred on an emission band of C_2 at 5,140 Å and with the same 39-arc s radius aperture. This aperture, which is the one we used most frequently, had a corresponding radius at the nucleus of the comet which ranged between 33×10^3 and 12×10^3 km during March and April 1986. To remove, at least to first order, the effects of the variation in aperture size at the comet, we have converted the observed C_2 fluxes to column densities and then to production rates assuming a Haser model. While one might quibble with the applicability of the Haser model to comet Halley, the results presented here concerning the short-period variability of the comet in no way depend on the strict validity of this model.

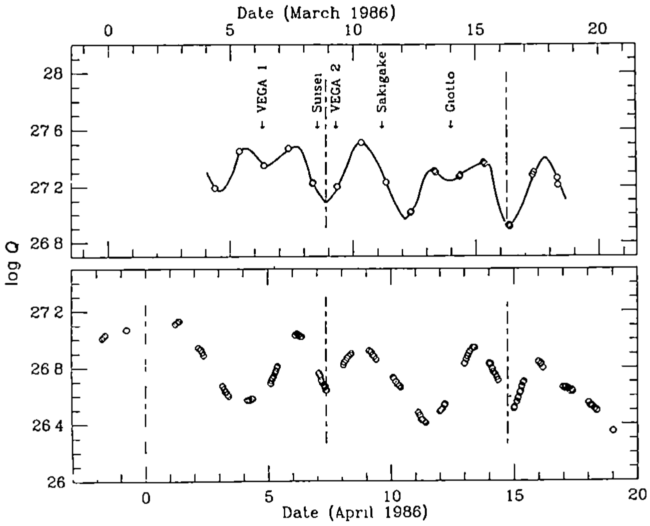


Fig. 2 Production rate of C_2 as a function of time in March and April 1986. Arrows mark the times of the various spacecraft encounters. Dashed lines identify corresponding minima in March-April

The observed production rates of C_2 in comet Halley during March and April are plotted as a function of time in Fig. 2. Because the comet was observable for most of each night in April, the continuity of the data from that month is sufficiently good that there is no question about the timescale of the variation. Furthermore, it is evident from even a casual inspection of Fig. 2 that the photometric behaviour of the comet between 3.5 and 11 April was repeated almost exactly between 11 and 18.5 April, indicating that the light curve has two maxima per cycle with a period near 7.5 days. During the March run, the period was evidently similar, but the light curve had a more complex shape than in April. Contemporaneous International Ultraviolet Explorer (IUE) observations with the FES (fine error sensor) camera[9] agree with the shape of the mean March light curve as we have drawn it through the observed points.

In addition to the periodic variation, a steady secular decrease in the production rate of C_2 can be seen in Fig. 2 throughout March and April. This secular change, which presumably resulted from the decreasing insolation as the comet moved away from the Sun, has been removed from the data before performing the period search described below.

A search for periodicity in the observations was performed by applying the method of phase dispersion minimization separately to the March and April data. This method[10] is ideal for application to data such as those in Fig. 2 because it makes no assumptions about the shape of the light curve. In this method, the data are phased according to an assumed period and then divided into a number of phase bins, *M*, spanning the full phase interval. The variance within each bin, $S_j^2(j=1, M)$, is computed, and then the overall variance for all bins S^2 is calculated from

$$S^2=\frac{\sum (n_j-1)S_j^2}{\sum n_j - M}$$

where n_j is the number of data points in the *j*th bin. A statistic, θ, defined as the ratio of S^2 to the variance of the total data set, is computed. This process is repeated for a broad range of trial periods. If the assumed period is very different from the true period, the phased data will constitute a scatter diagram and $\theta = 1$. At a true period, however, θ will reach a local minimum.

In Fig. 3*a*, we have plotted θ as a function of period for the April data. In computing θ, the data were divided into five phase bins. Note that over the entire range of periods from 1 to 10 days, the deepest minimum is at ~7.4 days. There is no significant minimum near 2.2 days. Figure 3*b*, *c* shows a plot of θ against period at higher resolution for the March and April observations. For the April data, a sharp minimum is found at 7.37±0.07 days. The quoted uncertainty is estimated. Because of less complete sampling of the light curve in March as compared with April, the minimum centred at 7.41 days in the θ versus period plot has a broad, flat bottom. The uncertainty in the period determination based on the March data alone is estimated to be ~0.1 day. Another minimum at 3.9 days was found in the analysis of the March observations. However, that period, which is almost exactly four times the sampling interval, can be shown to be an artefact of the period-finding method.

Composite light curves for March and April, constructed by phasing the data according to the 7.37-day period, are plotted in Fig. 4. We have adopted this single period for the two data sets because the two values determined above are the same to within their estimated errors and because we feel that the April run provided the more accurate value. The superposition of

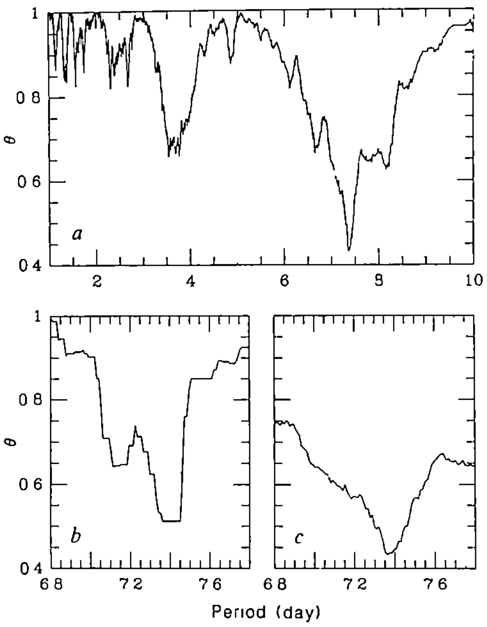


Fig 3 *a*, The statistic θ plotted as a function of period for the data between 29 March and 19 April At a true period, θ is expected to show a local minimum By far the deepest minimum seen here corresponds to a period near 7 4 days *b* And *c*, the same type of plot as described above shown at expanded time resolution for the 4-18 March and 29 March-19 April data, respectively

successive cycles is surprisingly good Certainly, one would not expect the light curve of an active comet to be as repeatable as that of an inert object like an asteroid, and, indeed, the form of comet Halley's light curve did evolve between the two observing runs The abscissae in Fig 2 have been located so that the March data are placed relative to the April data, as if shifted forward in time by four cycles of the 7 37-day period Note that the minima marked by the dashed lines nearly coincide in phase In fact, the primary change in the light curve between March and April is the absence in April of the first hump of the broad maximum seen in March A change of this character in the shape of the light curve could easily occur if a localized region of activity on the comet's nucleus became dormant or, at least, substantially decreased its output during the last 2 weeks in March The underlying periodicity of the light curve, on the other hand, would be unaffected

We believe that the most natural explanation of the periodic variations that we have observed in the brightness of comet Halley's inner coma is an up-and-down fluctuation in the comet's overall level of activity as surface regions of greater or lesser volatility are brought into sunlight by the rotation of the nucleus The observed period of the brightness variations is, therefore, identified with the synodic period of rotation of the cometary nucleus If this interpretation is correct, the observed period would be expected to have decreased by ~0 1 day between March and April if the sense of rotation is the same as that of the orbital motion Such a change is consistent with our observations However, our period determinations are not sufficiently accurate to unequivocally confirm a change of this magnitude

Comet Halley's light curve is seen to have basically two maxima and two minima per rotation While not required, such a light curve is not unexpected given the shape of the nucleus as revealed by the Vega and Giotto imaging experiments[11,12] Comet P/Arend-Rigaux, whose nucleus also has been shown to be elongated, has a two-peaked light curve[13]

Of the previous attempts to ascertain the rotational period of comet Halley, some, like the present study, were based on ground-based photometry, albeit with much less continuous observational coverage[6,7] The results have been inconclusive Whipple[5] and Sekanina and Larson[1,2] have taken a different approach based on the morphology of the coma as recorded primarily on ground-based photographs Sekanina and Larson find a rotational period near 53 h Investigators analysing observations from the Japanese ultraviolet imager experiment and the Vega 1 and 2 imaging experiments also have reported periods near that value[3,4] However, pending a fuller description of the spacecraft results, we are unconvinced that their period determination should be accepted as firm The ultraviolet imager on the Planet-A Suisei spacecraft detected several brief Ly α flashes over a period of months However, the total interval is extremely undersampled Continuous coverage has been published for only one 2 4-day interval, during which several flashes were seen If several flashes can occur per cycle and if the light curve has not been shown to repeat over even two successive cycles, then determination of a unique period would seem difficult The Vega 1 and 2 investigators have measured the rotational period of comet Halley's nucleus by examination of images obtained during two 40-min observation periods separated by 3 days While we in no way wish to denigrate the remarkable accomplishment that the Vega 1 and 2 images represent, they are of low resolution It seems questionable that a truly unique period determination could be based on those images

Given the strong periodicity at 7 4 days in the photometric data reported here and the absence in these observations of evidence for other periods, we believe that previous determinations of the rotational period of comet Halley should be re-examined We can think of no source of the variations present in our data other than rotation of the nucleus If this interpretation of the photometric data is correct, then we can predict the relative orientation of the nucleus at the times of the encounters of the comet with the Planet-A, Vega 1 and 2, and Giotto spacecraft The times of these encounters are indicated by arrows in Fig 2 We expect that at the Vega 1 and Giotto encounters the orientation of the nucleus with respect to the Sun was very similar At the time of the Vega 2 encounter, however, roughly the opposite side of the nucleus was facing sunward Suisei and Sakigake were not equipped to image the nucleus

We thank the staff of Cerro Tololo Inter-American Observatory for logistical and technical support CTIO is supported by NSF under contract no AST 78-27879 Ralph Martin assisted

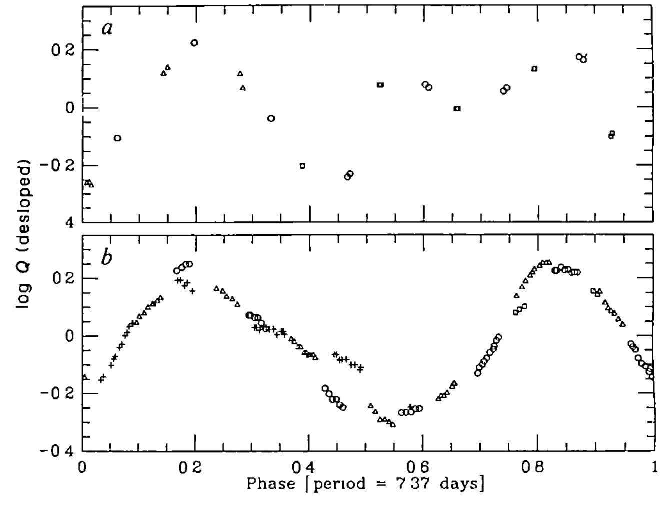


Fig 4 Composite light curves of comet Halley constructed by phasing the data in Fig 2 according to a 7 37-day period *a*, 4-18 March, *b*, 29 March-19 April The secular decrease in the mean brightness of the comet evident in Fig 2 has been removed before plotting Successive cycles of the variation are plotted with different symbols The different cycles are delineated by dashed lines in Fig 2 and are plotted here in order of increasing time as squares, circles, triangles, and pluses

with the observations in April, and Tobias Kriedl provided software for the period search, as well as helpful discussion of the phase dispersion minimization method. This research was supported by NASA grant NGR-03-003-001.

Received 17 September; accepted 31 October 1986.

1. Sekanina, Z. & Larson, S. M. *Nature* **321**, 357–361 (1986).
2. Sekanina, Z. & Larson, S. M. *Astr. J.* **92**, 462–482 (1986).
3. Kaneda, E., Ashihara, O., Shimizu, M., Takagi, M. & Hirao, K. *Nature* **321**, 297–299 (1986).
4. Sagadeev, R. Z. *et al. Nature* **321**, 259–262 (1986).
5. Whipple, F. L. in *Comets* (ed. Wilkening, L.) 227–250 (University of Arizona Press, Tucson, 1982).
6. Sekanina, Z. *Astr. Astrophys.* **148**, 299–308 (1985).
7. Meech, K. J., Jewitt, D. & Ricker, G. R. *Icarus* **66**, 561–574 (1986).
8. A'Hearn, M. F., in *Solar System Photometry* (ed. Genet, R.) 3-1-3-32 (Willmann-Bell, Richmond, 1983).
9. Festou, M. C. *et al. Nature* **321**, 361–363 (1986).
10. Stellingwerf, R. F. *Astrophys. J.* **224**, 953–960 (1978).
11. Sagadeev, R. Z. *et al. Nature* **321**, 262–266 (1986).
12. Keller, H. U. *et al. Nature* **321**, 320–326 (1986).
13. Millis, R. L., A'Hearn, M. F. & Campins, H. *Bull. Am. astr. Soc.* **17**, 688 (1985).

Cyanogen jets in comet Halley

Michael F. A'Hearn*, Susan Hoban*,
Peter V. Birch†, Craig Bowers†,
Ralph Martin† & Daniel A. Klinglesmith III‡

* Astronomy Program, University of Maryland, College Park, Maryland 20742, USA
† Perth Observatory, Bickley, Western Australia 6076, Australia
‡ NASA-Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

Current theory holds that jets, spirals and other detailed features in cometary comae are visible only in the continuum reflected by the dust particles. It has generally been thought that these features could not be seen in the emission bands of the typical neutral species, at least in part because the neutral parent molecules expand in directions perpendicular to any linear feature in a time short compared with that required to travel the length of the features. Consequently, the daughter species normally observable at optical and ultraviolet wavelengths are produced by dissociations that take place far from the point of origin of the gas. Here we present images of comet Halley taken in the light of CN emission, that show prominent jets extending to >60,000 km from the nucleus and lasting several weeks. We interpret these images as jets of submicrometre particles, perhaps the organic 'CHON' particles detected from spacecraft, from which the CN is directly produced.

During comet Halley's post-perihelion phase, we carried out observations with the 24-inch planetary patrol telescope at Perth Observatory, using a CCD (charge-coupled device) camera system which is one of two developed jointly with S. M. Larson (the system will be described elsewhere). The present observations were made in April, when the comet was near opposition, with an effective focal ratio of 5.4 and a scale of 1.8 arc s per pixel on the CCD chip. The filter set is a duplicate of the IAU standard filters distributed to the photometry and polarimetry net of the International Halley Watch and the photometric calibration was done with exactly the same techniques as used by that net.

The comet was imaged several times each night in various emission-band and continuum filters. Jets were visible in real time on the television monitor in the raw images taken with the CN filter. These features are not visible in nearly simultaneous images taken in the reflected continuum, nor are they coincident with the sharper, shorter features seen by shift-difference processing in nearly simultaneous broad-band images processed by S. M. Larson (personal communication). The jets seen in CN light appear to be of much higher contrast, broader, longer and in totally different orientations.

Figure 1 shows the processing that has been applied to all subsequent images. At the upper left is a direct image taken on 23 April, corrected by bias subtraction and flat field and smoothed with a three-point triangle (Hanning function). In all images presented here, north is at the bottom and east is at the left. This image, only a portion of the total frame from the CCD, is 5 arc min = 10^5 km square. The existence of three jets is clear: they extend roughly south, east and north-west but curve substantially into spirals. Images in the continuum have not yet been subtracted from our images in CN but photometric calibration indicates that the continuum will be negligible except in the few pixels nearest the nucleus. For processing, the image was extracted to an (r, θ) plot as shown at the lower left, in which radius is the ordinate and azimuth is the abscissa. To enhance the jets each horizontal line of the (r, θ) representation was renormalized to the maximum and minimum values along that line. The result is shown at the lower right, in which the shape of the jets is much more clearly seen than in the unnormalized representation. This plot was then converted back to the conventional image format at the upper right. This procedure was followed for all of the images shown in Fig. 2. All images were also scaled to the same physical size in kilometres at the comet.

It is clear that the jets have a spiral shape due to the rotation of the nucleus; despite the first impression, however, the view is not along the polar axis. Figure 2 shows a series of images from 16–23 April. During this period, the northern rotation pole (that is, the angular momentum vector) as defined at the 1910

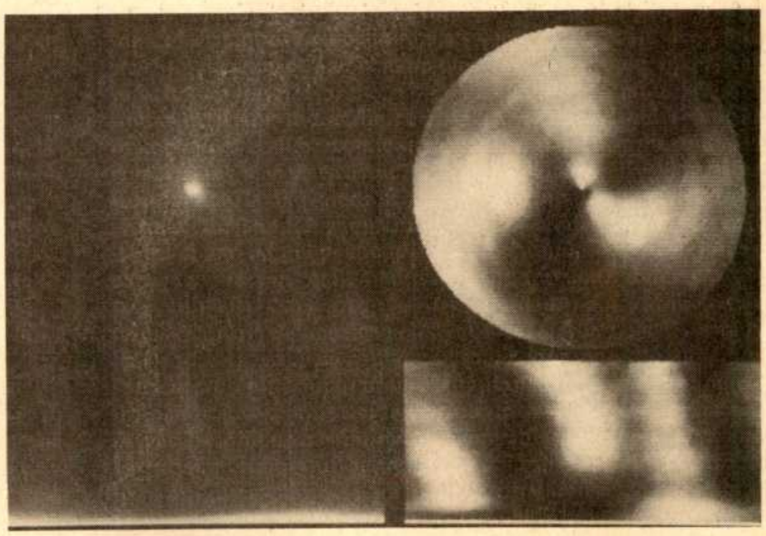


Fig. 1 Image processing steps applied to all images. The upper left panel is the cometary image extracted from the original CCD image, no. 81 in Table 1. Bias subtraction and flat-field corrections have been applied. The lower left panel is a radius (r) against polar angle (θ) display of the original image. The radial distance increases upward and the position angle on the sky is the abscissa varying from west through north. The lower right panel is a contrast enhancement of the lower left panel. The enhancement is such that each line, a circle in the original image, has been scaled between the minimum and maximum of the display. This effectively removes all radial variation. The upper right panel is the enhanced (r, θ) plot converted back to an (x, y) coordinate system.

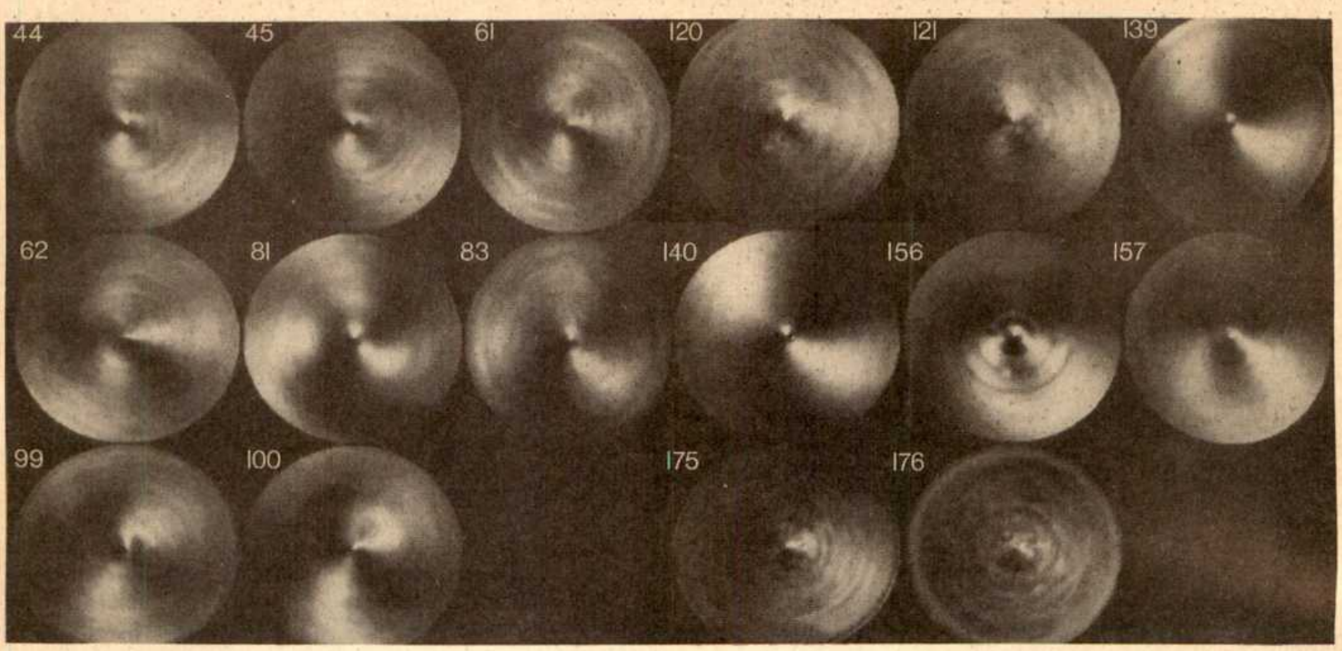


Fig. 2 The radially enhanced images for eight days at two frames per day. The details of the exposures are given in Table 1. Time increases with image number. Each image has undergone the enhancement process described in the text and Fig. 1.

apparition[1] was pointing nearly southward as seen on the sky and changed from nearly in the plane of the sky at the beginning of the period to pointing towards Earth, ~20° out of the plane of the sky. Preliminary indications from the present apparition are that the current rotation pole may differ by ~20° from that determined in 1910[2,3]. The position angle of the Sun varied across the entire south-west quadrant, from 199° on 16 April to 272° on 29 April.

Identification of jets persisting from one day to the next is dependent on the assumed period of rotation. Examination of the series of images, many of which are shown in Fig. 2, shows that there are at least three and, if the rotation period is 2.17 days[1], probably five jets involved. It is possible that other rotation periods may be consistent with the pattern, perhaps requiring fewer jets. A more detailed analysis will follow.

The ratio of the ejection velocity of the material in the jets to the rotational velocity of the nucleus can be estimated from the slope of the jets in the (r, θ) plots. With the assumptions that the spiral patterns are due to projected helices and that the rotational period is 2.17 days[1], the ejection velocity is $\sim\frac{2}{3}$ km s^{-1}, comparable to that inferred from spacecraft[4,5]. This scales inversely as the rotational period, but is only an order of magnitude estimate. Again, detailed modelling will follow.

The fact that the jets are so pronounced in the light of CN has significant implications for the formation of CN radicals. Figure 3 shows brightness profiles as a function of azimuth for several different radii in the image shown in Fig. 1: these represent horizontal cuts through the lower right panel and they show a range of 720° in azimuth (two full rotations) to exhibit more clearly features that would be at the edge of a plot showing only one rotation. The full width at half maximum (FWHM) of the CN distribution is ~45° at a radius of 30,000 km, which corresponds to a linear FWHM of 24,000 km. The emission from the jet must be due to fluorescence rather than, for example, prompt emission (emission immediately following dissociation into the excited $B^2\Sigma$ state) because a typical CN radical will fluoresce many times before leaving the jet, whereas it can produce only one photon by prompt emission. The half width at half maximum of the parent, then, must be $\leqslant$3,000 km, much too narrow for the parent to be a neutral gaseous species released from the nucleus. If we take the FWHM of the parent species to be 3,000 km at a radial distance of 30,000 km and take $\frac{2}{3}$ km s^{-1} as the ejection velocity, the maximum transverse velocity corresponds to a temperature near 5 K. In the absence of a confining mechanism for the jets, the inescapable conclusion is that the jets are composed of dust grains and that the CN is either dissociated directly from molecules on the surfaces of the grains or dissociated from parent molecules that have a very short ($\tau < 1{,}000$ s) lifetime before photodissociation and that vaporize from the grains. Although this result was quite surprising to us, Bobrovnikov[6,7] reported such jets in comets Halley and Westphal, and Wurm and Mammano[8] discussed the rapid formation of CN in expanding haloes.

Why do we not see these dust grains in the images taken in the continuum? As pointed out by Sekanina[3], the grains are probably no larger than 0.1 μm, so they would not be seen in the reflected continuum, which is dominated by much larger grains. In fact, the spacecraft that passed through the coma of Halley detected many more very small particles than had been expected[4,5,9,10,11]. It seems likely that the CN is being dissociated directly from the grains because estimates of the dissociation lifetimes of a variety of plausible parents for CN are all $>10^4$ s (ref. 12). It is tempting to suggest that the CN comes from the organic 'CHON' particles found by the spacecraft[13,14].

Table 1 Exposure data for selected CN imagery of comet Halley

Image no.	Date	Exposure		*R* (AU)	*Δ* (AU)
		Start (UT)	Duration (s)		
44	86.4.16	17:29:44	600	1.42	0.46
45		18:32:38	600		
61	86.4.17	14:05:46	600	1.43	0.48
62		16:01:04	900		
81	86.4.23	12:39:05	900	1.52	0.60
83		17:16:39	900		
99	86/4/24	13:23:51	900	1.54	0.63
100		14:53:00	1,200		
120	86/4/26	11:36:32	1,200	1.57	0.68
121		15:08:45	1,200		
139	86/4/27	15:25:37	1,200	1.58	0.71
140		17:46:04	1,200		
156	86/4/28	11:04:54	1,200	1.60	0.73
157		13:45:19	1,200		
175	06/4/29	11:52:06	1,200	1.61	0.76
176		14:25:59	600		

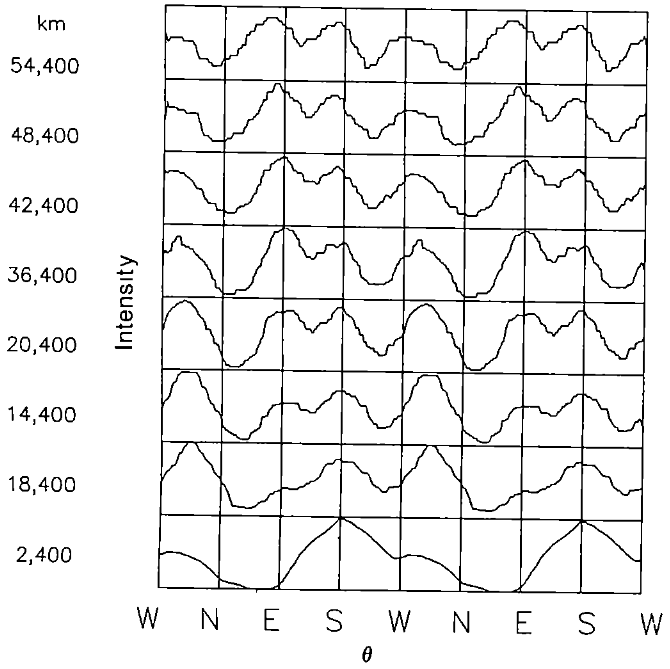


Fig 3 Plot of normalized intensity against azimuth for eight discrete radii projected on the sky, labelled by the value of projected radius These plots are horizontal cuts through the lower right panel of Fig 1, image no 81 in Table 1

A very crude estimate of the fraction of the CN in the comet that is produced in this way can be obtained by examination of the unnormalized version of Fig 3 At radii of 10,000–20,000 km, the fraction of the column density of CN that is in the jets is ~10–15% Because the high ejection velocity of CN on photodissociation causes the CN in the jets to diffuse rapidly into the background, it is not unreasonable to assume that this represents the total fraction of all CN that comes from the grains

It is not yet clear whether the CN jets discussed here and attributed to dissociation from grains are related to the fluctuations in CN detected from the spacecraft Levasseur-Regourd *et al.*[15] reported that an excess of CN (relative to a vectorial model) was detected near the nucleus by the Optical Probe Experiment on Giotto, they attributed this excess to prompt emission In addition to the excess emission that they reported near the nucleus, examination of their Fig 2 shows fine structure in the CN flux, particularly an enhancement near log $r = 3$ 6, which is clearly above the noise and might be due to passage of Giotto through jets such as those reported here Similarly, Moreels *et al*[16] have reported that the three-channel spectrometer on Vega 2 also detected an excess of CN near the nucleus The spacecraft data clearly merit a more detailed examination in the context of additional ground-based data Our own images from the time near the spacecraft encounters do not show any prominent jets in CN, but those images have not yet been examined in sufficient detail to rule out the presence of such jets It does seem at least plausible that the excess CN seen by the spacecraft experiments is provided by dissociation from the organic grains, as we have suggested for the CN in the jets reported here If our hypothesis is correct, similar structures should be visible in images taken in the light of other radicals composed of 'CHON', provided that they have sufficient contrast with the background Preliminary studies of images in the light of C_2 indicate that the jets are present but weaker, because of the strong contribution by the continuum which does not exhibit jets

M F A'H and S H. thank the staff of the Perth Observatory and the University of Western Australia for hospitality and assistance during the course of these observations, particularly Arie Verveer for his technical assistance, and the government of Western Australia for their support through these institutions We thank Zdenek Sekanina and Nalin Samarasinha for several helpful discussions regarding the geometry of the jets and Z S for pointing out that the dust grains must be very small, Steve Larson for comparing these CN jets with the ones used in modelling the nuclear rotation, and for designing and building the CCD system, many individuals for comments regarding the relationship to the organic grains, Jurgen Rahe for pointing out the earlier work on CN structures, and Barbara Boothman for programming the algorithms for the (r, θ) extraction and providing general assistance with the computing The CCD system was constructed with financial support from the International Halley Watch office at the Jet Propulsion Laboratory The observing programme was supported by NASA grant NSG 7322 and NSF grant INT 85-14707 to the University of Maryland

Note added in proof The existence of submicrometre 'CHON' particles in comets has recently been discussed by Greenberg[17] as a consequence of accretion of interstellar grains

Received 11 August, accepted 16 October 1986

1 Sekanina, Z & Larson, S M *Astr J* **92,** 462–482 (1986)
2 Sekanina, Z & Larson, S M *Nature* **321,** 357–361 (1986)
3 Sekanina, Z *Adv Space Res* **5,** 307–316
4 Krankowsky, D *et al. Nature* **321,** 326–329 (1986)
5 Vaisberg, O L *et al. Nature* **321,** 274–276 (1986)
6 Bobrovnikov, N T *Astrophys. J* **66,** 145–169 (1927)
7 Bobrovnikov, N T *Astrophys J* **66,** 439–464 (1927)
8 Wurm, K. & Mammano, A. *Icarus* **6,** 281–291 (1967)
9 Mazets, S P *et al. Nature* **321,** 276–278 (1986)
10 Simpson, J A. *et al. Nature* **321,** 278–280 (1986)
11 McDonnell, J A. M *et al. Nature* **321,** 338–341 (1986)
12 Bockelee-Morvan, D & Crovisier, J *Astr Astrophys.* **151,** 90–100 (1985)
13 Kissell, J *et al. Nature* **321,** 280–282 (1986)
14 Kissell, J *et al. Nature* **321,** 336–337 (1986)
15 Levasseur-Regourd, A. C *et al. Nature* **321,** 341–344 (1986)
16 Moreels, G *et al. Nature* **321,** 271–273 (1986)
17 Greenberg, J M ESA-SP **249,** 47–55 (1986)

Nitric acid cloud formation in the cold Antarctic stratosphere: a major cause for the springtime 'ozone hole'

Paul J. Crutzen* & Frank Arnold†

* Max Planck Institute for Chemistry, PO Box 3060,
D-6500 Mainz, FRG
† Max Planck Institute for Nuclear Physics, PO Box 103980,
D-6900 Heidelberg, FRG

Large depletions in stratospheric ozone were first reported by Farman *et al.*[1] at Halley Bay (76° S), and confirmed by satellite observations[2]. Chubachi[3] gives a detailed account of ozone decreases and temperatures in the lower stratosphere during the spring of 1982 at 69° S. There is now evidence[2] for annual declines in total ozone by ~6 and 3% in regions of total ozone minima and maxima, respectively, from September to mid-October since the late 1970s. We propose here a chemical mechanism for the formation of the ozone hole. It involves removal of gaseous odd nitrogen by ion- and/or aerosol-catalysed conversion of N_2O_5 and $ClONO_2$ to HNO_3 vapour, followed by heteromolecular HNO_3–H_2O condensation, leading to HNO_3–H_2O aerosols. At an altitude of 17 km, these processes start at temperatures below 205±5 K, well above the condensation temperature of pure water vapour. We propose that the absence of gaseous odd nitrogen and catalytic methane oxidation reactions driven by sunlight in early spring lead to large OH concentrations which rapidly convert HCl to ClOX. Catalytic reactions of ClOX and BrOX cause drastic ozone destructions and can account for the springtime 'ozone hole' first observed by Farman *et al.*[1]. By our model the depletion would be

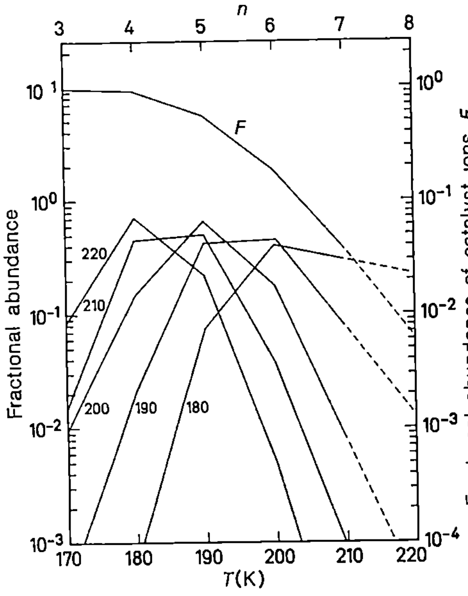


Fig. 1 Size distributions of $H^+(H_2O)_n$ cluster ions (upper abscissa and left ordinate) for an altitude of 17 km, an H_2O-vapour volume mixing ratio of 3×10^{-6} and temperatures of 180, 190, 200, 210 and 220 K (thermodynamical data of Lau *et al.*[17]) Also shown is the fraction *F* of $H^+(H_2O)_n$ ions (lower abscissa and right ordinate) with $n\geq6$ that may serve as catalysts

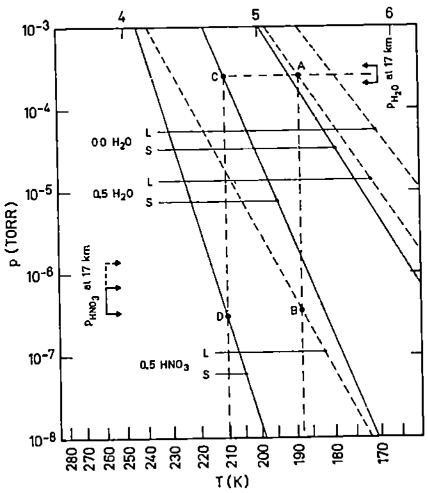


Fig. 2 Equilibrium saturation vapour pressures of H_2O and HNO_3 over solid (solid curves) and liquid (broken curves) H_2O-HNO_3 mixtures with HNO_3 mass fractions of 0.5 and 0.0 (pure H_2O) Also shown are partial pressures for H_2O and HNO_3 as observed in the background atmosphere around 17 km altitude[29] Condensation temperatures for liquid mixture (points A and B) and solid (points C and D) mixtures are about 188 K and 211 K respectively For liquid and solid H_2O ($f_2=0.0$) condensation temperatures are about 179 K and 191 K respectively

mainly due to emissions of industrial organic chlorine compounds. Arctic regions may also become affected. The depletion lasts while HNO_3, but not HCl, is incorporated in the particles in the temperature range 205±5 K to 192 K.

Farman *et al*[1] connected the ozone decrease with the extremely cold temperatures and stable dynamics of the lower stratosphere and also proposed that the depletion was due to the increasing transfer of industrial chlorine compounds to the stratosphere This idea was later developed by McElroy *et al*[4] and Solomon *et al*[5] McElroy *et al*[4] consider reactions between ClO and BrO Their scheme requires the absence of NO and NO_2 from the stratosphere, brought about by reactions of N_2O_5 and $ClONO_2$ on the surfaces of the stratospheric cloud particles observed[6,7] in polar regions

$$N_2O_5+H_2O\rightarrow2HNO_3,\ ClONO_2+H_2O\rightarrow ClOH+HNO_3 \quad (1,2)$$

and furthermore the conversion of most HBr to BrO_x, and most HCl to ClO_x For the latter assumption, however, they do not present a mechanism Solomon *et al*[5] assume that efficient reactions of $ClONO_2$ and HCl occur in polar clouds

$$ClONO_2+HCl\rightarrow Cl_2+HNO_3 \quad (3)$$

leading to the catalytic reaction chain

$$OH+O_3\rightarrow HO_2+O_2,\ Cl+O_3\rightarrow ClO+O_2 \quad (4,5)$$

$$ClO+HO_2\rightarrow ClOH+O_2,\ ClOH+h\nu\rightarrow Cl+OH \quad (6,7)$$

$$\text{net} \quad 2O_3\rightarrow3O_2 \quad (D1)$$

The efficiency of the heterogeneous reactions (1)-(3) for stratospheric conditions has not been measured, so the above studies are speculative

To explain the observations we present a chemical mechanism, involving the following steps (i) conversion of NO_x to N_2O_5 during polar night

$$NO+O_3\rightarrow NO_2+O_2,\ NO_2+O_3\rightarrow NO_3+O_2 \quad (8,9)$$

$$NO_3+NO_2(+M)\rightarrow N_2O_5(+M) \quad (10)$$

(ii) formation of HNO_3 by ion- and /or aerosol-catalysed conversion of N_2O_5 and $ClONO_2$ at temperatures ≤215-210 K, (iii) heterogeneous, heteromolecular condensation of HNO_3 vapour with H_2O giving rise to stratospheric cloud particles and depletion of gaseous HNO_3 at temperatures ≤205±5 K, (iv) efficient hydroxyl radical formation by cosmic ray ionization and ozone photolysis, amplified by photochemical methane oxidation reactions after sunlight returns to the polar regions, (v) conversion of HCl to ClOX, and HBr to BrOX, by reaction with OH, (vi) ozone destruction through ClOX and BrOX catalysis

Conversion of N_2O_5 and $ClONO_2$ to HNO_3 and the depletion of gaseous NO_x and HNO_3, which are possible only at low temperatures, are key elements in our mechanism as they prevent

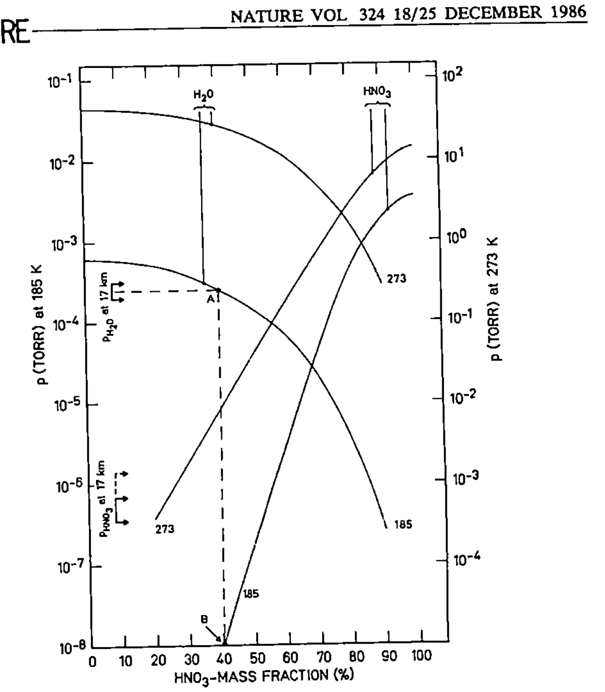


Fig 3 Equilibrium saturation vapour pressures of H_2O and HNO_3 over liquid H_2O-HNO_3 mixtures against the HNO_3-mass fraction f_2 for temperatures of 273 K and 185 K Also shown are partial pressures of H_2O and HNO_3 as observed in the background atmosphere around 17 km altitude Both sets of curves show the lowering of equilibrium vapour pressures of H_2O and HNO_3 over those of the one-component systems At 185 K and $f_2=40\%$ large supersaturation of gaseous HNO_3 occurs and co-condensation of HNO_3 and H_2O takes place

OH removal by the otherwise most important catalytic reaction cycle

$$OH + HNO_3 \rightarrow H_2O + NO_3, \quad NO_3 + h\nu \rightarrow NO_2 + O \qquad (11, 12)$$

$$NO_2 + OH(+M) \rightarrow HNO_3(+M) \qquad (13)$$

$$\text{net} \qquad 2OH \rightarrow H_2O + O \qquad (H1)$$

Conversion of N_2O_5 and $ClONO_2$ to HNO_3 via (1) and (2) may be catalysed by ions and/or aerosols. The ion-catalysed mechanism involves an H_2O molecule as a reaction partner, contained in a cluster ion[8-10]. Atmospheric ambient ions present at ~17 km, the centre of the height region of interest, are $H^+(H_2O)_n$, $H^+(CH_3CN)_m(H_2O)_n$, and $NO_3^-(HNO_3)_m(H_2O)_n$ as was found from recent balloon-borne mass-spectrometer measurements[11-13]. For $H^+(H_2O)_n$ ions, for example, N_2O_5 conversion may proceed via

$$N_2O_5 + H^+(H_2O)_n \rightarrow HNO_3 + H^+(H_2O)_{n-1}HNO_3 \qquad (1a)$$

$$H^+(H_2O)_{n-1}HNO_3 + H_2O \rightarrow HNO_3 + H^+(H_2O)_n \qquad (1b)$$

becoming exothermic only for $n \geq 6$. Ions of the type $NO_3^-(HNO_3)_m(H_2O)_n$ may be even more efficient catalysts as extra energy is gained due to stronger clustering of HNO_3 formed by the N_2O_5 reaction[10]. However, to serve as catalysts both positive and negative ions must be hydrated enough. This is possible only at low atmospheric temperatures. Fig 1 shows calculated relative abundances of the members of the $H^+(H_2O)_n$ family and also the fraction F of $H^+(H_2O)_n$ ions with $n \geq 6$. Taking the calculated F values, lower limits to the N_2O_5 lifetime, t_i, against ion-catalysed conversion can be estimated. For a total ion concentration of ~10^4 cm^{-3} and a typical rate coefficient for an ion-molecule reaction of 2×10^{-9} cm^3 s^{-1} one obtains $t_i \geq 5 \times 10^4/F$ s. Hence, at temperatures below about 215 K, t_i may become shorter than one month. It should be mentioned, however, that exothermicity does not necessarily guarantee a large reaction coefficient and t_i may therefore be larger than this. For a more detailed discussion of ion processes see ref 10.

Conversion of N_2O_5 and $ClONO_2$ on aerosols is rather uncertain as the reaction probability per collision, γ, is not well known. The lifetime t_{AC} of a gaseous molecule against collision with aerosols in the background atmosphere at ~17 km is ~10^4 s and thus the reaction time t_A becomes t_{AC}/γ. To make t_A one month, $\gamma = 2 \times 10^{-3}$ is required, which is much larger than γ measured[14] for concentrated sulphuric acid at 300 K.

In the cold Antarctic stratosphere, however, the aerosol content may become much larger than in the background stratosphere as found from solar extinction measurements[6,7]. This excess aerosol, termed polar stratospheric clouds (PSCs) has previously[7] been explained in terms of H_2O condensation alone. Here we shall argue that heteromolecular condensation of HNO_3 and H_2O occurs and leads to a great increase of the aerosol mass and surface area, which, in turn, may accelerate aerosol-catalysed N_2O_5 and $ClONO_2$ conversion, and more importantly leads to a depletion of HNO_3 vapour. Figure 2 shows extrapolations of measured[15] equilibrium saturation vapour pressures e_1 (H_2O) and e_2 (HNO_3) over an HNO_3-H_2O mixture with an HNO_3-mass fraction $f_2 = 0.5$. The extrapolations contain considerable uncertainties, especially for the solid state[10]. The case $f_2 = 0.5$ was chosen as this mixture gives the highest condensation temperature. Also shown in Fig 2 are partial pressures for a_1 (H_2O) and a_2 (HNO_3) as found in the background atmosphere around 17 km altitude[29]. The procedure to derive such conditions is illustrated in Fig 3 for liquid mixtures and a temperature of 195 K. The condition for heteromolecular condensation ($e_1 \leq a_1$ and $e_2 \leq a_2$) for $T =$ 185 K is met at f_2 near 40%. For this mixture $e_1 = a_1$, but $a_2 \gg e_2$,

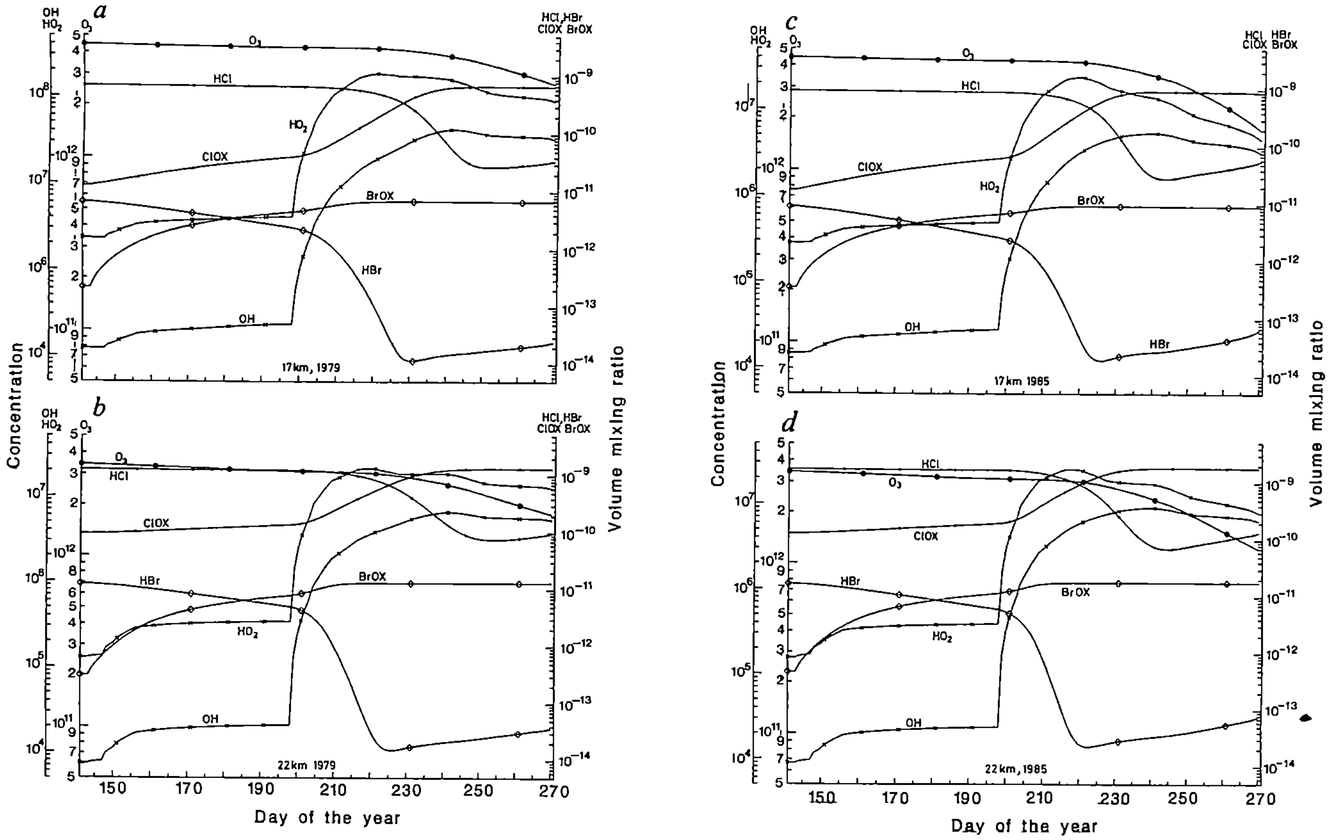


Fig. 4a, b, calculated concentrations (molecules cm^{-3}) at noon, of OH and HO_2 (first scale on the left) and of O_3 (second scale on the left), and volume mixing ratios of HCl, ClOX, HBr and BrOX (right scale). Polar night conditions at 70° S, here defined as Sun below horizon for 24 h period, lasted from 22 May (day 142) to 18 July (day 198). It was assumed that from day 147 nitric acid became incorporated in the polar stratospheric clouds. Calculations were performed for two altitudes, 17 km (a) and 22 km (b), and estimated conditions for the year 1979. c, d, Similar calculations as a, b, for 1985.

so that gaseous HNO_3 condenses out. This lowers both e_1 and the atmospheric partial pressure of HNO_3 and increases e_2, so that about 1.5 times more H_2O than HNO_3 molecules condense out. Because $a_1 \gg a_2$, the resulting atmospheric pressures of HNO_3 will be much reduced, while those of H_2O are hardly affected.

An inspection of Fig. 2 reveals that the condensation temperature T_c for a liquid mixture with $f_2 = 0.5$ (points A and B) is about 188 K. For solid mixture one obtains $T_c = 211$ K (points C and D). A further cooling below T_c by only 3-4 K would be sufficient to deplete 50% of the HNO_3 vapour, giving rise to an increase of the aerosol mass and surface area by factors of about 100 and 20 respectively. Condensation of pure H_2O vapour, which becomes possible only at temperatures below about 191 K (solid) to 179 K (liquid), gives rise to an additional strong increase of the aerosol mass. The heteromolecular HNO_3-H_2O condensation mechanism proposed here offers an explanation for recent extinction measurements[6,7] showing a strong increase of the stratospheric total extinction at 1 μm wavelength over Antarctica as temperatures fall below about 205 K, which is close enough to 211 K considering the uncertainty of vapour pressure extrapolations and height distribution of the aerosol. Below about 191 K also other trace gases, particularly HCl, may become incorporated in aerosols[10]. With the increased aerosol surface, aerosol-catalysed N_2O_5 and $ClONO_2$ conversion may become important, depending on γ. However, ion-catalysed processes may be more efficient. It is possible that they are the missing source of HNO_3 in the polar winter stratosphere required in recent studies[10,18].

Another ion process which has a role in polar-night stratospheric chemistry is hydroxyl radical formation by ion chemistry driven by galactic cosmic rays. The rate of this ion-assisted OH formation is one to two OH molecules per ionization event or ~40-80 OH molecules $cm^{-3}\,s^{-1}$ at high geomagnetic latitudes. Because the galactic cosmic ray ionization rate varies with solar activity, an 11-yr modulation will be induced. A more detailed discussion of ion and aerosol processes will be given in ref. 10.

Modelling of stratospheric photochemical processes is complex and all details cannot be discussed here. Our model uses a standard set of chemical reactions[20,21]. To emphasize the effects of chemistry alone, we adopted a very small eddy diffusion coefficient, because the south polar vortex seems to be such a stable dynamic system[22].

When gas-phase NO_x and HNO_3 are removed from the lower stratosphere, the OH radial reacts mainly with CH_4, HCl and HBr

$$CH_4 + OH(+O_2) \rightarrow CH_3O_2 + H_2O, \; HCl + OH \rightarrow Cl + H_2O \qquad (14, 15)$$

$$HBr + OH \rightarrow Br + H_2O \qquad (16)$$

The results of the model calculations shown in Fig. 4, for altitudes of 17 and 22 km for 1979 and 1985, at 70° S, are mainly intended to demonstrate the sensitivity of ozone to odd-nitrogen removal during springtime. They are not close simulations of the exact conditions in the Antarctic. The calculations were started with concentrations obtained with our two-dimensional model. Ozone and temperature data were taken from the measurements[3] made at Syowa (60° S) in 1983. The small jumps in the curves occurring around day 147 indicate the date at which we have assumed that HNO_3 becomes incorporated in the cloud particles. The large changes in concentrations which occur near day 200, about 20 July, mark the return of daytime conditions (Sun above the horizon) for some period during the day. Results are shown for noon. In these calculations ClOX denotes the sum of Cl, ClO, ClOH, Cl_2O_2 and OClO. The last three compounds all dissociate rapidly in daylight, producing ClO_x (= Cl + ClO). BrOX is defined similarly. The stratospheric concentrations of total inorganic bromine were assumed to be 1% of those of inorganic chlorine, which seems to be roughly in agreement with the observations, although the uncertainty is large[23,24].

Besides reaction cycle D1, there are two other catalytic reaction chains that are important in the breakdown of stratospheric ozone during early spring when sunlight has returned to the Antarctic.

$$ClO + BrO \rightarrow Cl + Br + O_2, \; Cl + O_3 \rightarrow ClO + O_2 \qquad (17, 5)$$

$$Br + O_3 \rightarrow BrO + O_2 \qquad (18)$$

$$\text{net} \qquad 2O_3 \rightarrow 3O_2 \qquad (D2)$$

as considered by McElroy *et al.*[4], and

$$ClO + ClO(+M) \rightarrow Cl_2O_2(+M) \qquad (19)$$

$$Cl_2O_2 + h\nu \rightarrow Cl + ClOO, \; \lambda < 350\,\text{nm} \qquad (20)$$

$$ClOO + M \rightarrow Cl + O_2 + M \qquad (21)$$

$$Cl + O_3 \rightarrow ClO + O_2 (2\times) \qquad (5)$$

$$\text{net} \qquad 2O_3 \rightarrow 3O_2 \qquad (D3)$$

as recently introduced by Molina and Molina[25] and R. A. Cox (personal communication). According to Molina and Molina, the photolysis rate of Cl_2O_2 is roughly that of ClOH. The product yield of (20) is still only based on indirect evidence[25]. For (19) we chose the rate expression $k = 6 \times 10^{-33}\,(300/T)^{2.1}\,cm^6\,molecule^{-2}\,s^{-1}$ by R. A. Cox (personal communication).

We note from the results in Fig. 4 that during polar night much HBr can be converted to BrOX by (16). The conversion would be much enhanced if HBr reacted rapidly with NO_3, which seems quite likely. The conversion of HCl to ClOX is less efficient, because (15) is ~1/25 as fast as (16). During polar night, a large fraction of the OH radicals produced by the galactic cosmic rays is lost by reaction with CH_4. The situation changes drastically after sunlight returns, when, starting from low odd hydrogen concentrations, reactions with CH_4 lead to rapid net production of odd hydrogen. The important reactions are

$$CH_4 + OH(+O_2) \rightarrow CH_3O_2 + H_2O, \qquad (14)$$

$$CH_3O_2 + HO_2 \rightarrow CH_3O_2H + O_2 \qquad (22)$$

$$CH_3O_2H + h\nu \rightarrow CH_3O + OH, \; \lambda \leq 350\,\text{nm} \qquad (23)$$

$$CH_3O + O_2 \rightarrow CH_2O + HO_2 \qquad (24)$$

$$\text{net} \qquad CH_4 + O_2 \rightarrow CH_2O + H_2O$$

followed with about 30% probability by

$$CH_2O + h\nu \rightarrow H + CHO, \; \lambda \leq 330\,\text{nm}, \qquad (25)$$

$$H + O_2 + M \rightarrow HO_2 + M \qquad (26)$$

$$CHO + O_2 \rightarrow CO + HO_2 \qquad (27)$$

The net result of the oxidation of one methane molecule by OH is therefore an average net gain of 0.6 odd hydrogen radicals, providing for a rapid buildup of OH radicals until these are also significantly lost by

$$CH_3O_2H + OH \rightarrow CH_3O_2 + H_2O, \; CH_3O_2H + OH \rightarrow CH_2O + H_2O + OH \qquad (28a, b)$$

For the first weeks, the main primary source of OH is production by galactic cosmic rays. Later, the reactions

$$O_3 + h\nu \rightarrow O(^1D) + O_2, \; \lambda \leq 310\,\text{nm}, \qquad (29)$$

$$O(^1D) + H_2O \rightarrow 2OH \qquad (30)$$

become more important, further enhancing the OH concentrations. ClOX is efficiently produced from HCl after the polar night, leading to efficient destruction of ozone mainly through reaction cycle D1 and D3, with some contributions also from the BrO_x/ClO_x cycle D2. The calculations for two altitudes 17 and 22 km, shown in Fig. 4, demonstrate the sensitivity of stratospheric ozone to the perturbations that characterize the springtime Antarctic lower stratosphere, so that the observed trends in ozone removal can be explained by the chemical mechanisms presented above The ozone decay continues as long as most HNO_3 is bound in the stratospheric polar clouds, that is, until the Antarctic stratospheric warming occurs in late September or early October[3,22] Then the evaporation of the polar clouds will reintroduce the various NOX gases into the lower stratosphere through known photochemical processes and most ClOX will be converted back to HCl The decreasing ozone concentrations will delay the springtime solar warming of the lower stratosphere and this could gradually increase the period of unstable photochemistry The ozone depletion will last as long as HNO_3, but not HCl is incorporated in PSCs At an altitude of 17 km this occurs between 205±5 K and 192 K Although less frequent than in the Antarctic, such conditions occur also in the Arctic stratosphere

In this paper we could not present a detailed analysis of all features of the Antarctic ozone hole, because this requires consideration of and access to extensive meteorological, chemical and optical data sets[26] We have, however, clearly shown how in the Antarctic lower stratosphere during winter and springtime unstable photochemical conditions can develop that lead to drastic destruction of ozone Several other complications that are likewise not treated in this paper must be considered in future studies, such as the effects of the major eruptions of volcano El Chichon in 1982, the possibility of gravitational settling of the particles[26], and the reactions of odd hydrogen and other chemical compounds on the particles

To test our chemical mechanism, measurements of odd nitrogen and chlorine gases as well as ambient ions would be very informative No measurements of HNO_3 and HCl have so far been reported in the Antarctic for September–October Available total column NO_2 observations[27,28] agree with the mechanism proposed, but do not provide proof Also laboratory studies of the binary H_2O-HNO_3 system at low temperatures and ion-catalysed reactions are needed

We thank Christoph Bruhl and Barnabas Thwaites for their assistance with the calculations, and Ian Galbally for valuable comments

Received 8 September, accepted 11 November 1986

1 Farman, J C, Gardiner, B G & Shanklin, J D *Nature* **315,** 207–210 (1985)
2 Stolarski, R. S *et al. Nature* **322,** 808–811
3 Chubachi, S in *Handbook for MAP* Vol 18 (ed Kato, S) 453–457 (SCOSTEP, Univ Illinois, Urbana, 1985)
4 McElroy, M B, Salawitch, R. J, Wofsy, S C & Logan, J A *Nature* **321,** 759–762 (1986)
5 Solomon, S, Garcia, R. R, Rowland, F S & Wuebbles, D J *Nature* **321,** 755–758 (1986)
6 McCormick, M P, Steele, H M, Hamill, P, Chu, W P & Swissler, T J *J atmos Sci.* **39,** 1387–1397 (1982)
7 Steele, H M, Hamill, P, McCormick, M P & Swissler, T J *J atmos Sci.* **40,** 2055–2067 (1983)
8 Ferguson, E E, Fehsenfeld, F C & Albritton, D L in *Gas Phase Ion Chemistry* Vol 1, 45–81 (Academic Press, 1979)
9 Bohringer, H, Fahey, D W, Fehsenfeld, F C & Ferguson, E E *Planet. Space Sci.* **31,** 185–191 (1983)
10 Arnold, F & Crutzen, P J (in preparation)
11 Arnold, F in *Proc. DAHLEM Workshop on Atmospheric Chemistry* (Ed Goldberg, E D) 273 (1982)
12 Arnold, F & Buhrke, T *Nature* **301,** 293 (1983)
13 Ferguson, E E & Arnold, F *Acct. chem. Res* **14,** 327 (1981)
14 Baldwin, A C & Golden, D M *Science* **206,** 562–563 (1979)
15 Clavelin, J L. & Mirabel, P *J Chim. phys* **76,** 533–537 (1979)
16 Forsythe, W R. & Giauque, F *J Am. chem. Soc.* **64,** 48–61 (1942)
17 Lau, Y K., Ikuta, S & Kebarle, P *J Am. chem. Soc.* **104,** 1462–1469 (1982)
18 Austin, J A, Garcia, R R, Russell, J M, Solomon, S & Tuck, A F *J geophys Res* **90,** 5477–5485 (1986)
19 Evans, W J F, McElroy, C T & Galbally, J E *Geophys Res Lett* **12,** 825–828 (1985)
20 DeMore, W B *et al. Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling,* JPL Publication 85-37 (NASA Jet Propulsion Laboratory, California, 1985)
21 Crutzen, P J & Schmailzl, U *Planet Space Sci.* **31,** 1009–1032 (1983)
23 Labitzke, K. *J geophys Res* **86,** 9665–9678 (1981)
23 Berg, W W, Crutzen, P J, Grahek, F E, Gitlin, S N & Sedlacek, W A *Geophys Res Lett.* **7,** 937–940 (1980)
24 Berg, W W, Heidt, L E, Pollock, W, Sperry, P D & Cicerone, R J *Geophys Res Lett* **11,** 429–432 (1984)
25 Molina, L T & Molina, M *J phys Chem.* (submitted)
26 McCormick, M P & Trepte, C R. *J geophys Res* (submitted)
27 McKenzie, R. L & Johnston, P V *Geophys Res Lett.* **11,** 73–75 (1984)
28 Shibasaki, K. in *Handbook for MAP* Vol 18 (ed Kato, S) 506–509 (SCOSTER, Univ Illinois, 1985)
29 World Meteorological Organization *Atmospheric Ozone 1985,* WMO Rep no 16 (WMO, Geneva, 1985)

Are domains in liquid crystalline polymers arrays of disclinations?

Barbara A. Wood & Edwin L. Thomas†

* Polymer Products Department, E I DuPont, Experimental Station, Wilmington, Delaware 19898, USA
† Polymer Science and Engineering Department,
University of Massachusetts, Amherst, Massachusetts 01003, USA

Interaction of mobile disclinations in the liquid crystalline state results in various disclination arrays. The mesophase state of liquid crystalline polymer (LCPs) samples has been assumed to consist of a polydomain texture, each domain having an internal structure of uniform orientation bounded by regions where the molecular director field undergoes reorientation from one domain to its neighbour. However, the detailed nature of the boundary region (in particular whether the molecular director reorients in a continuous or discontinuous fashion) is largely unknown. We have observed groupings of disclinations of alternating sign which resemble Bouligand's serrated wall structure[1] and 'foursomes' of two positive and two negative disclinations which bear a likeness to the 'effective domain' structure recently hypothesized by Marrucci[2] to account for certain rheological aspects of LCPs. The arrangements of disclinations in our thin-film samples give rise to an apparent polydomain texture consisting of regions of uniform molecular orientation bounded by disclinations where the molecular director undergoes continuous rotation about an axis parallel to the disclination lines (Neel walls) The specific disclination arrays observed will be influenced by the molecular weight and stiffness constants of the particular LCP and by the flow history, thickness and boundary conditions of the film[3]; nevertheless, except in the vicinity of the disclination cores, the field of molecular director trajectories throughout the polydomain sample texture is both continuous and translationally parallel.

Due to industrial and academic interest in stiff chain and semiflexible polymer molecules, there have been an increasing number of structural and rheological studies of liquid crystalline polymers in recent years A banded texture is often noted in flow-oriented LCP specimens Donald *et al*[4] have examined the banded structures in detail and have recently shown that annealing transforms the banded morphology into a 'vein' structure which they relate to orientation domains Initial explanations for phenomena such as the yield stress and the three-region viscosity-shear rate curve (two shear-thinning regions separated by a newtonian plateau) have invoked the polydomain concept where, during flow, the domains slide by one another easily at low shear rates At moderate shear rates, the domains are supposed to break up, leaving dispersed domains in matrix of uniform molecular orientation, and, at high shear rates, a continuous monodomain structure is realized[5] A recent theory of Marrucci[2], however, emphasizes the role of disclinations in the rheology of liquid crystalline polymers The fluid is imagined

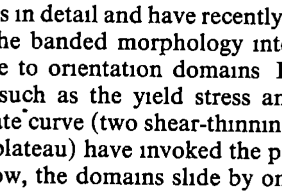


Fig 1 Schematic of an effective domain in a liquid crystalline polymer consisting of an array of four half-integer disclinations Lines indicate the molecular trajectory about the defect lines (●) (from ref 2)

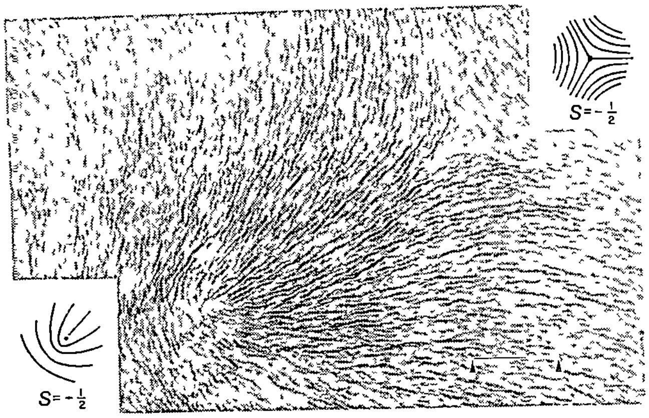


Fig 2 TEM image of disclination dipole in a thermotropic liquid crystalline polymer and schematic (insets) of molecular director trajectories around a positive and a negative half-integer disclination Scale bar, 10 μm

to consist of an assembly of 'effective domains' due to a particular arrangement of disclinations In two dimensions, Marrucci's scheme consists of a group of two positive and two negative half-integer disclinations (Fig 1) This particular array yields an approximately zero average net director orientation over the basic domain and hence zero global sample orientation Moreover, the molecular director changes continuously over the polydomain sample except at the disclination cores Wissbrun[6] has further developed Marrucci's model to allow the size (or number density) of the domains to change during flow, with consequent changes in the rheological characteristics Recent flow visualization experiments using light microscopy have suggested that the flow can both increase[7] and decrease[8] the number of defects in the fluid To date, no detailed experimental observations of the proposed disclination-domain structures have been made

Recent transmission electron microscopy (TEM) studies of alternating rigid/flexible thermotropic LCPs conducted in our laboratory have permitted direct visualization of the molecular director pattern in flow-oriented thin films at a resolution unmatched by previous light micrographs[9] The mesophase texture and defects existing in the polymer melt state are retained by thermal quenching of the polymer fluid to room temperature By annealing the glassy liquid crystalline polymer film above its glass transition but below the melting point, crystalline lamellae grow perpendicular to the local chain axis and effectively 'decorate' the molecular director field throughout the sample

Samples were prepared by annealing of substrate-free oriented thin films of the thermotropic polyester from 1,10-decane *bis*(terephthaloyl chloride) and hydroquinone[9,10] Figure 2 is a TEM image of a region containing two singularities in the field of oriented long, thin crystalline polymer lamellae Electron diffraction proves that the chain axis is normal to the long dimension of the lamellae Thus, the set of lamellar normals represents the set of molecular director trajectories over the sample As was recently pointed out by Frank[11], the lamellar decoration pattern demonstrates that the molecular director field consists of translationally parallel trajectories, that is, families of congruent curves such that the vector joining points of the same slope on any two members of the family is a constant The trajectories of the lamellar normals in the vicinity of each defect corresponds quite well with that expected from the molecular director trajectories associated with 180° disclination lines running perpendicular to the specimen plane The molecular director trajectories about such isolated defects are in good accord with those predicted nearly 30 years ago by Frank[12] Dipoles consisting of a positive and negative 180° disclination at a typical separation of micrometres are frequently observed Apparently, annihilation occurs if pairs of opposite sign disclinations approach each other more closely The core regions of the negative (threefold symmetric) disclination appear bright, which indicates a local thinning of the film, possibly corresponding to a high concentration of chain ends and hairpin molecular folds

In addition to isolated and variously paired disclination dipoles, arrays involving approximately equal numbers of positive and negative disclinations are frequently observed A diamond-shaped group involving two positive and two negative half-integer disclinations have been seen repeatedly (Fig 3*a*) Details of the molecular director trajectories are shown in Fig 3*b* The molecular director pattern is nearly identical to that of the hypothetical "effective domain" structure envisaged by Marrucci (see Fig 1) The groups of four disclinations observed also resemble the 'lozenges' defined by edge dislocations of layers in cholesteric small-molecule liquid crystals as described by Bouligand[1] Groups of two positive and two negative disclinations (although of integer strength) were first drawn by Lehmann in a 1918 monograph[13]

A larger group of disclinations is shown in Fig 4*a* A several square micrometres region of uniform orientation is bounded by a serrated wall of 10 alternating positive and negative disclinations Bouligand's[1] schematic of a serrated wall structure is shown in Fig 4*b* The discrete appearance of domains suggested by the images shown in Fig 3*a* and 4*a* results from observing the field of normals to the molecular director trajectories rather than the trajectories themselves[11] Regions of approximately

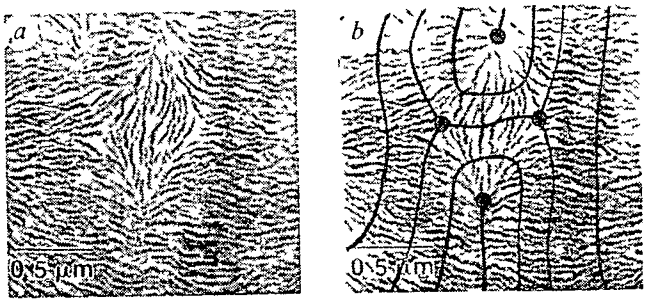


Fig 3 *a*, TEM image and *b*, schematic overlay of lamellar trajectories in a diamond-shaped region defined by two pairs of oppositely signed half integer disclinations Scale bars, 0 5 μm

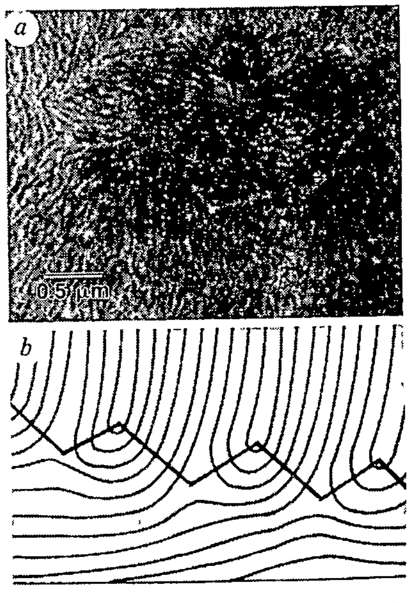


Fig 4 *a*, TEM image of region of approximately uniform orientation defined by walls of alternating positive and negative disclinations and *b*, schematic of a disclination wall (ref 1) Scale bar, 0 5 μm

uniform orientation are enclosed by disclinations where continuous rotation of the molecular director occurs about an axis parallel to the disclination lines (Neel walls in the language of magnetism) Except in the vicinity of the disclination cores, the field of molecular director trajectories is both continuous and translationally parallel

We thank the NSF Materials Research Laboratory of the University of Massachusetts for financial support and the Monsanto Company for a fellowship for B A W

Received 27 May, accepted 14 October 1986

1 Bouligand, Y *J Phys* **34**, 604-614 (1973)
2 Marrucci, G *Proc IX Int Congr Rheol* (University of Mexico, 1984)
3 Kleman, M *JCS Faraday Discuss* **79**, 215-224 (1985)
4 Donald A M , Viney, C & Windle, A. H *Phil Mag* **B52**, 925-941 (1985)
5 Asada, T & Onogi, S *Polym. Engng Rev* **3**, 324-353 (1983)
6 Wissbrun, K *JCS Faraday Discuss* **79**, 161-173 (1985)
7 Graziano, D J & Mackley,M *Molec Cryst Liq Cryst* **106**, 73-93 (1984)
8 Asada, T , Onogi S & Yamase, H, *Polym Engng Sci* **24**, 355-360 (1984)
9 Thomas, E L , & Wood, B A *JCS Faraday Discuss* **79**, 229-239 (1985)
10 Wood, B A thesis, Univ Massachusetts-Amherst (1985)
11 Frank, F C *JCS Faraday Discuss* **79**, 274-280 (1985)
12 Frank, F C *JCS Faraday Discuss* **25**, 19-28 (1958)
13 Lehmann, O *Die Lehre von den flussigen Krystallen* (Bergman, Wiesbaden 1918)

Rapid reduction of nitrogen oxides in exhaust gas streams

R. A Perry & D. L. Siebers

Combustion Research Facility, Sandia National Laboratories,
Livermore, California 94550, USA

Nitrogen oxides (NO_x) play a major role in the formation of photochemical smog and in acid rain production[1] Some progress has been made in reducing NO_x emissions through the use of combustion and exhaust control schemes, including three-way catalyst, staged combustion, and ammonia injection. Nevertheless, estimates indicate that the production of nitrogen oxides will continue to increase until the end of the century and beyond, if current trends continue[2] We describe here a new chemical process capable of completely removing NO_x from the products of combustion The effectiveness of this process is established in flow tube experiments. Further, we demonstrate the practical feasibility of this method by eliminating NO_x from a portion of the exhaust from a single-cylinder diesel engine. Based on these results we conclude that this process for rapid reduction of nitrogen oxides could play a major role in controlling NO_x emissions from most combustion devices.

Nitrogen oxide formation in combustion has received much attention[3-7] The recognition[4] that ammonia can be used to reduce nitric oxide in exhaust gases led to major efforts to characterize reactions important in the Thermal $DeNO_x$ process The mechanisms and models that resulted have greatly improved understanding of nitrogen chemistry in combustion For example, this research led to the prediction of nitric oxide concentrations in ammonia injection schemes[5,7] and in general nitrogen chemistry in combustion[6,7] Experimental efforts have continued to increase our understanding of the radical species important in the conversion of fuel-bound nitrogen to N_2 and NO during combustion Perry[8] studied reactions of the NCO radical with combustion species NCO radicals formed from the photolysis of isocyanic acid (HNCO) react very rapidly with nitric oxide, suggesting[8] that this reaction might be important in nitrous oxide formation in combustion These observations, combined with previous insights into nitrogen chemistry, led to a new scheme for NO_x reduction[9]

The proposed nitrogen oxide reduction scheme is based on the addition of isocyanic acid to an exhaust stream containing NO_x Isocyanic acid is formed from the decomposition of cyanuric acid (a nontoxic, commercially available compound) when cyanuric acid is heated above approximately 330 °C

$$(HOCN)_3 \xrightarrow{\Delta} 3HNCO$$

When HNCO is mixed with an exhaust gas stream at temperatures $\geq$400 °C, a series of reactions occur that result in the loss of HNCO and NO Using a long-pass optical cell with Fourier transform infrared (FTIR) detection, carbon monoxide, carbon dioxide, water, and nitrogen (by inference) are observed as products Table I shows results obtained in initial flow reactor experiments to study the overall chemistry of the proposed scheme This process has been used to remove more than 99% of the nitric oxide emissions in simulated and real exhaust gas streams containing from 250-3,000 p p m by volume of NO

The flow reactor was a 3 8 cm diameter stainless steel tube 1 m long, enclosed in a furnace The simulated exhaust gas was sampled with a 6 4 mm quartz probe, the position of the probe tip was varied over a range of 70 cm by the use of a stepping motor The flow rates of the gases used (Ar/NO/HNCO) were controlled and measured with calibrated mass-flow controllers Residence times in the flow reactor were determined from the equation $t = Ad/V$ where t is time, V is flow rate, A is the cross sectional area of the tube, and d is the distance from the point of mixing The temperature of the tube was measured using chromel/alumel thermocouples, and the pressure in the tube was measured using a capacitance manometer

Although the initiation chemistry is uncertain, some observations regarding the overall chemical process can be made from the flow-tube experiments The most prominent feature of the results (Table 1) is the decrease in NO and HNCO concentration and the formation of CO and CO_2 It is evident from the data that the process occurs on a 1 1 molar basis

Table 1 Flow data obtained from stainless steel flow reactor using simulated exhaust gases (Ar/NO/HNCO mixture)

Temperature (°C)	Pressure (torr)	$[NO]_i$ (p p m)	$[HNCO]_i$* (p p m)	$[NO]_f$† (p p m)	[CO]‡ (p p m)	$[CO_2]$‡ (p p m)	$[HNCO)_f$‡ (p p m)	Time (s)
585	508	0	660	0	420	190	50	0 30
582	512	740	610	410	30	500	80	0 30
639	540	725	695	120	100	520	75	0 30
638	539	500	635	70	130	425	80	0 59
647	637	370	350	100	40	250	60	0 15
647	537	365	485	95	80	345	60	0 72
640	537	725	720	80	130	525	65	1 15
643	532	725	805	45	160	600	60	1 40
663	537	670	925	40	80	800	20	2 00
651	603	530	555	115	65	450	40	0 35
644	537	725	855	25	210	600	45	1 70
643	537	725	820	45	160	600	60	1 40

* $[HNCO]_i$ is the initial HNCO concentrations calculated assuming carbon mass balance from the equation $[HNCO]_i = [HNCO]_f + [CO_2] + [CO]$
† [NO] determined by NO_x chemiluminescent analyser
‡ [CO], $[CO_2]$ and [HNCO] measured using absorption at selected lines precalibrated with known concentrations of CO, CO_2, HNCO

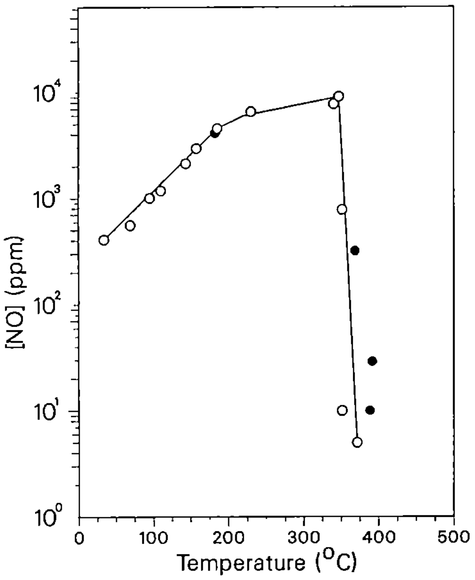


Fig 1 Plot of NO concentration (p p m) in treated portion (1 0 1 min^{-1}) of diesel engine exhaust gas versus cyanuric acid bed temperature (°C) The diesel engine was operated at medium load (overall equivalence ratio of 0 5) and 1,800 r p m Instrumentation sensitivity is limited to ~5 p p m Data were obtained on successive days, with solid circles taken initially and open circles obtained in separate experiments

The following possible reaction scheme, when combined with previously developed models[5–7], may explain these experimental results

$$\begin{aligned} HNCO &\xrightarrow{\Delta} NH+CO \\ NH+NO &\rightarrow H+N_2O \\ H+HNCO &\rightarrow NH_2+CO \\ OH+CO &\rightarrow H+CO_2 \\ NH_2+NO &\rightarrow N_2H+OH \\ &\rightarrow N_2+H_2O \\ N_2H &\rightarrow N_2+H \end{aligned}$$

This reaction scheme implies that, following initiation, the sequence through N_2H produces a growth of one hydrogen atom per cycle to sustain the chain One important aspect of the chemistry of our proposed method is the absence of the need for controlled amounts of oxygen, suggesting that it could be readily implemented as a control process, if the presence of oxygen does not greatly change the mechanism

To test the effect of excess oxygen on the process, and to develop a technology for exhaust conditions that are presently untreatable by existing technologies, such as 3-way catalyst and ammonia injection, we used the proposed method to reduce the nitrogen oxides in a portion of a diesel engine exhaust gas stream The experiments used a single cylinder, 7 2 horse power Onan diesel engine The engine was operated at a medium load condition (an overall equivalence ratio of 0 5) and an engine speed of 1800 r p m A continuous sample of unfiltered exhaust gas was drawn at a rate of 1 l min^{-1} and passed over cyanuric acid in a 14 6 cm long by 4 72 cm diameter stainless steel tube with a volume of 256 cm^3 (the cyanuric acid bed) The sample then flowed through a 17 8 cm long by 2 2 cm diameter tube packed with two sizes of stainless steel spheres, diameter 0 55 cm and 0 64 cm (the packed bed), which had an open volume of 31 cm^3 and a total surface area including steel spheres of 1,630 cm^2 The packed bed was used to test the effect of surface-to-volume ratio and residence times on the chemistry Finally, the exhaust gas flowed through a NO_x chemiluminescence analyser Samples also were taken of the treated and untreated exhaust gas for FTIR analysis

Figure 1 shows the results of the diesel engine experiments The NO concentration measured in the treated exhaust gas is plotted as a function of the cyanuric acid bed temperature For cyanuric acid bed temperatures above 360 °C the NO_x was essentially eliminated These results were insensitive to packed bed temperatures between 450 °C and 900 °C, provided that the cyanuric acid bed temperature was greater than 360 °C At lower acid bed temperatures, nitric oxide levels in excess of that produced by the diesel engine (410 p p m) were observed (Fig 1) Surface interactions of HNCO with oxygen are believed to be important in this effect, and suggest the need to minimize the surface area in designing a practical device Experiments using gaseous HNCO injection in oxidizing environments, which show that NO-reduction takes place efficiently in the presence of excess oxygen, will be reported with related results in detail

In summary, we have discovered a chemical means of removal of NO_x from the products of combustion or industrial emissions, and we have demonstrated its feasibility on a portion of the exhaust from a diesel engine This process has important implications for a wide spectrum of energy conversion technologies and for industrial applications that presently produce NO_x as exhaust byproducts

We thank Dan Ferko, Jim Boehmke, and Mark Jaska for technical support Supported by US Department of Energy, Office of Basic Energy Sciences, Division of Chemical Physics, and Office of Energy Utilization Research, Division of Energy Conversion and Utilization Technologies

Received 8 October, accepted 31 October 1986

1 Finlayson-Pitts, B J & Pitts, J N Jr, *Atmospheric Chemistry – Fundamentals and Experimental Techniques* (Wiley, New York, 1986)
2 *Acid Rain Information Book Final Report DE81 024267* (US Department of Commerce National Technical Information Service, Washington, 1981)
3 Miller, J A., Smooke, M D , Green, R. M & Kee, R J *Combust Sci. Technol* **34**, 149 (1983)
4 Lyon, R K. US Patent No 3,900 554, August 1975
5 Miller, J A., Branch, M C & Kee, R. J *Combust Flame* **43**, 81–98 (1981)
6 Miller, J A *et al. 20th Int Symp*, 673–684 (Combustion Institute, 1984)
7 Hanson, R. K. & Salimian, S *Combustion Chemistry* (ed Gardiner, W C Jr, 361 (Springer, New York, 1984)
8 Perry, R A *J Chem. Phys* **82**, 5485–5488 (1985)
9 Perry, R A US Patent pending

A magmatic heat pump

S. A. Morse

Department of Geology and Geography,
University of Massachusetts, Amherst, Massachusetts 01003, USA

The structures of igneous cumulates can provide valuable clues to inaccessible magmatic processes. Many mafic layers in the Rhum Intrusion have 'finger structures'[1–3] that protrude into the overlying felsic rocks. These structures are attributable to melting of the overlying olivine–feldspar rocks by hot mafic magma emplaced as sills[4]. The bases of mafic layers contain smooth-bottomed olivine cumulates. Here I show that the heat needed to melt the roof of the layer is carried by rapidly rising light solute released by growth of basal olivine crystals that were deposited by two phase-convection. Compositional convection therefore drives a heat pump transferring heat of fusion from remote growing crystals to those that are melting. This action contrasts with cooling under a refractory roof, which is conduction-limited. The heat pumping principle shows why ultramafic magmas may melt their roofs whereas felsic magmas tend to form upper border zones.

Details of the crystallization sequence in the finger structures reported at Rhum have led to the realization that the mafic layers are pene-contemporaneous sills whose presence has caused melt-corrosion of the overlying rock[4] Details are clarified in ref 4, but the features central to the present discussion may be summarized as follows Finger structures occur only in troctolite (olivine-plagioclase rock, often called allivalite) overlying mafic layers rich in olivine cumulate They are interpreted as melting features because the host rock fabric is continuous across the fingers, which therefore represent passive removal of felsic material and subsequent crystallization of mafic material The crystallization sequence olivine-pyroxene in the fingers shows that the mafic material was supercooled when it finally solidified in the finger spaces[4] The initial mafic magma was hot enough to dissolve the troctolite, so it was in the olivine field, not yet co-saturated with plagioclase feldspar and certainly not with pyroxene The mafic layers have thicknesses commonly in the range 5–100 cm

Melt features seem to be absent at the bases of the mafic layers (There are load structures into crystal mushes of troctolite[3], but these do not require melting at the local site) The magma layers were emplaced into essentially solid troctolite, mineral compositions above and below the layers are similar, typically $Fo_{84\pm2}$ in olivine, for example[5] The initial state immediately after injection was therefore that of a hot, essentially isothermal magma layer sandwiched between cooler felsic rocks Melting should have begun immediately at both boundaries Instead, there is abundant evidence of melting at the upper boundary, but no firm evidence of melting at the lower

The first obvious difference between the top and bottom boundaries is the presence of a basal olivine cumulate layer below (Fig 1) This crystal layer can act as a thermal blanket and as a physical barrier to percolation of hot, dense liquid Three kinds of convection are relevant two-phase, compositional, and thermal The rapid deposition of olivine crystals is promoted by two-phase convection[6,7] of crystal-liquid suspensions from the cooler top of the mafic magma layer (This result differs from those based on calculations of the Rayleigh number that suggest enduring turbulence and suspension instead[8] However, the latter were reached without consideration of two-phase convection An alternative route to a quite different set of Rayleigh numbers has been suggested by Marsh[9]) Growth of olivine releases a light solute[10] that convects freely upward[11], it also locks the local temperature onto a thermal maximum called a 'latent heat hump'[12] The resulting compositional convection is an important agent in causing solidification of the porous olivine cumulate[10,11,13,14] The rapid formation of such a solid caprock is required to prevent percolation of the magma to the underlying layer The entire convective cycle is driven by absorption of heat at the roof of the mafic magma layer, which, in turn, drives two-phase convection , which carries downwards the heat sink needed to induce growth at the floor and thereby to induce compositional convection

Thermal convection runs in the same direction as compositional convection but is less important A useful measure of the relative potentials of compositional verus thermal convection is provided by the density ratio $R_\rho = \beta\Delta S/\alpha\Delta T$, where $\beta\Delta S$ is a compositional expansion times a concentration change and $\alpha\Delta T$ is a thermal expansion times a temperature change The density ratio has been evaluated for the growth of cumulates[15], and found to lie between six and 10^6, with the larger values more common Such large values show that the compositional contribution to buoyancy greatly exceeds the thermal contribution, and hence that the convection is dominantly compositional

Corrosion of the roof rocks is extensive (Fig 1) The latent heat released by the crystallization of 1 g of olivine can melt 2 g of troctolite[16] Accounting for the lower density of the troctolite, a given volume of dunite can account for removal of ~2 5 times as much troctolite The olivine present in the rest of the mafic layer, along with the basal olivine, can easily account for the missing troctolite in terms of latent heat The thermal budget of a crystallizing magma sheet melting its roof was addressed by Irvine[17] Applying the relations of his Fig 14 to the present case, it can be judged that about twice as much cumulate can form when the heat evolved is absorbed by melting at the roof, compared to conduction alone through the roof

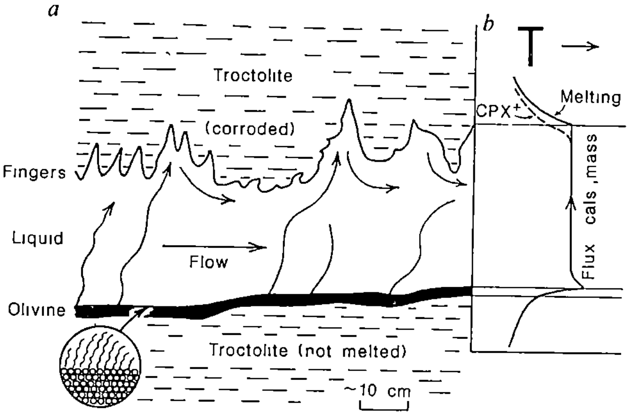


Fig 1 *a*, Rhum finger structures above a mafic layer with an olivine cumulate base (black), after a photograph (Fig 163 of ref 2) The cotectic felsic+mafic material (troctolite) below the mafic layer shows no evidence of melting, in contrast to that in the base of the overlying layer This asymmetry is accounted for by strong compositional convection of light rejected solute away from the growing basal olivine layer (shown by the magnified view in the lower circle) This convection pumps the heat of fusion through the mafic liquid to melt the overlying troctolite, where heat is absorbed The thermal relations are shown schematically in *b* When the convectional stirring becomes weak, supercooling within the fingers allows metastable growth of pyroxene (dashed curve labelled CPX⁺) The thermal maximum (latent-heat pump) at the top of the olivine cumulate arises from the local maximum of latent heat released[12] A small supercooling of the average magma is assumed, to allow for nucleation of olivine The mafic magma may be 100 °C or so hotter than the troctolite

It is proposed, therefore, that recrystallization of olivine in the mafic magma released most of the heat needed to dissolve the roof rocks Rising streams of rejected solute (Fig 1) will stir the magma layer and carry heat of crystallization away from the growing interface[12] Complementary two-phase plumes of olivine-seeded magma will carry supercooling downwards to allow growth of the basal cumulate, which is aided by conduction through the floor The resulting temperature profile is portrayed in Fig 1 The latent-heat hump occurs where the most olivine is growing, near the cumulate interface The average magma temperature is shown as supercooled to allow for the nucleation and growth of olivine and eventual metastable crystallization of pyroxene Physical stirring by the rising rejected solute keeps the liquid essentially isothermal The ambient temperature at the roof is high enough to melt the roof rocks The compositional convection therefore acts as a heat pump from the crystallizing base of the layer to the melting roof The melted material is flushed away from the propagating fingers and dissolved in the feldspar-undersaturated mafic magma The process must be rapid to avoid heating the walls of the fingers excessively, because they ultimately chill the liquid into the metastable pyroxene field[4] The chilling presumably occurs when the heat-pumping process has died down below some critical threshold velocity

The magmatic heat pump is similar to a classical isothermal heat pipe driven by the counterflow of two phases such as water and steam The magma layer is also sensibly isothermal, but here the counterflow consists of a two-phase flow originating at one interface balanced by a single-phase, density contrast flow originating at the other The pump and pipe are similar in transferring latent heat from one boundary to another

Melting on a much grander scale evidently occurred at the Muskox Intrusion in Canada, where the roof rocks of the intrusion were extensively melted[17] during the crystallization of the olivine-rich cumulates It seems that heat pumping by compositional convection occurred here on a massive scale. The paired heat source (olivine cumulate) and sink (fusible roof rocks) were both present, and compositional convection must have operated strongly. The pumping action could be damped but not erased by any convective stirring, because compositional convection will act over at least part of the distance of travel. If paired two-phase and compositional convective streamers or plumes dominated the convective flow pattern, heat pumping may have acted directly from floor to roof.

By contrast, felsic cumulates have a dense rejected solute[10] that cannot transport heat upwards through the overlying magma. Roof melting would therefore not be expected in felsic intrusions and frozen upper border zones are observed at such felsic intrusions as Skaergaard[15] and Kiglapait[10]. The rollover point between dense and buoyant rejected solute occurs, in mafic magma series, approximately at magnetite saturation[15], hence magnetite-bearing calc-alkaline magmas can more readily melt their roofs by heat pumping from cumulates than tholeiitic and more reduced magmas. This principle accords with the common impression that calc-alkaline series magmas tend to show more contamination effects than Fenner-trend magmas. The extended calc-alkaline effect is much weaker than the effect seen in very mafic magmas, because the latent heat of fusion of magnetite is less than that of magnesian olivine and magnetite generally crystallizes in company with feldspar, so that the bulk contributions of both latent heat and buoyant solute are small compared with magmas crystallizing olivine.

Although roof melting is not anticipated for felsic intrusions, the drainage of hot, dense solute from felsic roof cumulates implies that floor melting might occur by heat pumping. This could be the case for certain anorthosites and other felsic cumulates early in their history. However, this process is self-limiting because it produces an upper thermal blanket of roof cumulates through which conduction will have to occur for the magma to crystallize. By contrast, heat pumping with roof melting continually exposes cooler material to the magma, thus maximizes cooling and crystallization rates[17].

Heat pumping by compositional convection may account for roof melting on scales ranging from individual layers at Rhum to whole intrusions emplaced into fusible crust. Pumping is most efficient when the basal cumulate is ultramafic and the roof is readily melted. This combination may occur most commonly for early olivine cumulates, but might also occur for later magnetite cumulates. The crystallization of magnetite in company with other mafic crystals and feldspar renders most calc-alkaline cumulates weakly capable of pumping. Inverted heat pumps from felsic roof rocks are inherently self-limiting because they impose an insulting cumulate in the path of heat removal to the Earth's surface.

Support of NSF grant 8408320 is acknowledged.

Received 11 August; accepted 14 October 1986.

1. Brown, G. M. *Phil. Trans. R. Soc.* **B240**, 1-53 (1956).
2. Wager, L. R. & Brown, G. M. *Layered Igneous Rocks* (Freeman, San Francisco 1967; Oliver & Boyd, London, 1968).
3. Butcher, A. R., Young, I. M. & Faithfull, J. W. *Geol. Mag.* **122**, 491-502 (1985).
4. Morse, S. A., Owens, B. E. & Burcher, A. R. *Geol. Mag.* (in the press).
5. Butcher, A. R. *Geol. Mag.* **122**, 503-518 (1985).
6. Grout, F. F. *J. geol.* **26**, 481-499 (1918).
7. Morse, S. A. *Geol. Mag.* **123**, 205-214 (1986).
8. Huppert, H. E. & Sparks, R. S. J. *Contr. Miner. Petrol.* **75**, 279-289 (1980).
9. Marsh, B. D. *Abstr. Prog. geol. Soc. Am.* **17**, 653 (1985).
10. Morse, S. A. *Geol. Soc. Am. Mem.* **112**, 204p (1969).
11. Tait, S. R., Huppert, H. E. & Sparks, R. S. J. *Lithos* **17**, 139-146 (1984).
12. Morse, S. A. *J. geophys. Res.* **87**, A10-A18 (1982).
13. Bowen, N. L. *Am. J. Sci.* **39**, 175-191 (1915).
14. Walker, D., Jurewicz, S. & Watson, E. B. *Eos* **66**, 362 (1985).
15. Morse, S. A. *J. Petrol.* **27** (in the press).
16. Ghiorso, M. S. & Carmichael, I. S. E. *Contr. Miner. Petrol.* **71**, 323-342 (1980).
17. Irvine, T. N. *Can. J. Earth Sci.* **7**, 1031-1061 (1970).
18. Wager, L. R. & Deer, W. A. *Meddr. Grønl.* **105**, 1-352 (1939).

Formaldehyde oxime leaching of metals from deep-sea manganese nodules and other ores

R. Derie

Department of Ore Dressing, Université Libre de Bruxelles,
50 Avenue Roosevelt, B-1050 Brussels, Belgium

Formaldehyde oxime, $CH_2{=}NOH$, has for a long time been recognized as a sensitive colorimetric reagent for manganese, copper and metals of the iron group. In alkaline medium, the oxime dissociates, forming the $CH_2{=}NO^-$ anion which bonds to multivalent metals, giving complexes of the general formula $M(CH_2NO)_6^{(6-n)-}$, where n is the oxidation state of the metal[1,2]. Here I report the results of experiments which show that aqueous formaldehyde oxime constitutes a new and effective leaching agent for several types of ores, such as pyrolusite, malachite, heterogenite and deep-sea manganese nodules. The reagent seems particularly attractive for nodules, as the leaching is rapid, in acid as well as in neutral and/or alkaline media, and at ambient and lower temperatures. Manganese nodules constitute an important potential source of nickel and especially cobalt: the amount of cobalt recoverable from the north-west equatorial Pacific alone is estimated[5] to be 4×10^6 tonne—of the same order as the world total land-based reserves.

The formaldehyde oxime solutions were prepared by mixing solid hydroxylamine hydrochloride with an equimolar amount of a 38 wt% aqueous solution of formaldehyde. The reaction is immediate, yielding a concentrated solution of formaldehyde oxime, which is then diluted and brought to the desired *p*H by addition of ammonia or sodium hydroxide.

The exact origin of the nodules, collected during the 1974–78 campaign of Ocean Mining Associates, is unknown, but their composition (Mn 28.9, Fe 5.02, Ni 1.29, Cu 1.04, Co 0.24, Zn 0.13, Mo 0.065, V 0.051 wt% on dry basis) is very similar to that of the nodules from the Clarion-Clipperton zone (eastern Pacific Ocean)[2]. The X-ray diffraction spectrum (Fe Kα line) consisted only of the four major, broadened, lines of todorokite, plus the main quartz line.

Other materials selected for the leaching experiments were a phosphorus-rich manganese ore (Mn 31.6, Ca 13.0, P 0.25 wt%) from Raouisse (Vanuatu) in which we identified pyrolusite (β-MnO_2) and calcite, and a complex Zairian copper-cobalt ore, consisting mainly of badly crystallized heterogenite (CoOOH) and malachite, together with some muscovite and diaspore.

All these materials were ground to <106 μm. The dissolution experiments were carried out in a thermostated vessel, fitted with a simple magnetic stirrer. At intervals, samples were taken from the reaction medium, filtered and then analysed by atomic absorption or atomic emission (DCP) spectrophotometry.

Pyrolusite (as well as laboratory-grade MnO_2) dissolves easily in formaldehyde oxime solutions. From *p*H 5-11 (experiments beyond this limit were avoided as the oxime is sensitive to oxidation in alkaline medium) a brownish-red coloration appears instantaneously, evidence of the formation of the $Mn(CH_2NO)_6^{2-}$ anion. The reaction is rapid: for example, at 20 °C and pH 9.8, with 1.5 g ore in 300 ml 1M formaldehyde oxime, >90% of the Mn passed into solution after 5 min, the leaching being complete after ~10 min. Under these conditions, the gangue minerals (calcite and apatite were identified in the residue) were left practically unaltered.

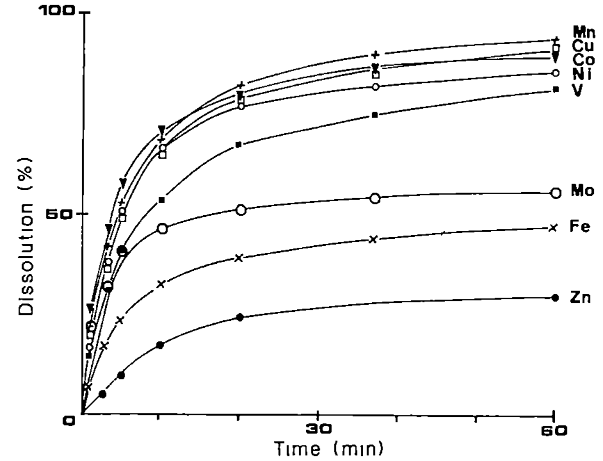


Fig. 1 The extraction of various elements from eastern Pacific manganese nodules (for conditions see text).

Separate dissolution experiments made under the same conditions as above have shown that among calcite, dolomite, magnesite, fluoroapatite and goethite, which are main gangue minerals of manganese ores, only calcite dissolved to a significant extent (25 mg Ca per litre solution after 60 min).

The quantity of manganese dioxide that can be dissolved in formaldehyde oxime solutions corresponds roughly to the stoichiometry of the reaction, that is, the formation of the hexacoordinated complex. Thus 1 litre of 1 M formaldehyde oxime solution dissolves 0.14 mol MnO_2 (12 g) at 20 °C and *p*H 9.8.

When the dissolution is done in an acidic medium (*p*H <5), the oxime is oxidized by MnO_2, and colourless solutions of Mn(II) are produced.

The powdered nodules dissolve even faster than pyrolusite in formaldehyde oxime; the solutions are also very dark. To slow the reaction rate the dissolution experiments were done in an ice bath.

Figure 1 shows an example of the results. The conditions were: temperature 0 °C, *p*H 7.5, 1.5 g nodule in 500 ml 0.5 M formaldehyde oxime. The percentage of every element passing into solution is represented as a function of time.

Other experiments were made under different *p*H conditions: as a rule, the dissolution paths of Mn, Ni, Cu and Co are always very similar (these elements are closely associated in the todorokite structure) and are not influenced by the alkalinity of the medium, in the *p*H range ~6–11.

The amount of iron dissolved never exceeded 70% of the iron content of the nodules, even if the leaching experiment was prolonged to 24 h at ordinary temperature. The mineralogical support of this insoluble iron was not identified, as only quartz, labradorite—a felspar-type mineral—and traces of a 7-Å clay mineral were identified by X-ray diffraction in the dissolution residue.

The dissolution behaviour of the other valuable elements of the nodules is more or less pH-dependent.

Table 1 Dissolution kinetics of copper and cobalt from a complex heterogenite—malachite ore (for conditions see text)

Time (min)	Co dissolved, wt %	Cu dissolved, wt %
1	12	5
3	27	9
5	43	12
10	64	16
20	70	21
30	77	27
75	94	34
120	98	40

By changing the experimental conditions, and particularly by raising the nodule: oxime ratio, we leached most (75–95%) of the nickel, copper and cobalt, leaving 45% of the manganese and >99% of the iron in the solid residue. These results could have a long-term economic interest, and will be discussed later (R.D., in preparation).

The third substance studied was the complex copper–cobalt ore. We observed that heterogenite always dissolves faster than malachite, and to ensure complete extraction of copper in a reasonable time it is necessary to operate at ≥35 °C. We did not succeed in leaching copper or cobalt selectively. As an example, table 1 gives the percentages of these elements dissolved, as a function of time at 20 °C and *p*H 7.0, with 1.5 g ore in 500 ml 0.5 M formaldehyde oxime.

Because most of the oxide, hydroxide and carbonate ores of copper and cobalt are easily leached by ammonia and/or mineral acids, economic interest in leaching by formaldehyde oxime seems unlikely. It is nevertheless remarkable that this leaching may be done in a neutral medium.

Little is known about the formation constants of metal–formaldehyde-oxime complexes. The only published data are those of Bečka and Jokl[4]; the values are high (log $K_{\text{form}} = 7.5$, 9.8, 17.3, and 18.9 respectively for the Cu (II), Mn (IV), Fe (III) and Ni (IV) complexes), indicating a stability much greater than, say, that of the ammine complexes of Cu and Ni. This alone cannot account, however, for the reactivity of formaldehyde oxime solutions towards the particular minerals studied here, as indicated by the fact that manganese dioxide is easily attacked by the oxime, but not iron oxyhydroxide (goethite), whereas the stability of the Fe (III) complex is seven orders of magnitude greater than that of the Mn (IV) complex.

The imperfections in the crystal lattices of the minerals must play a prominent role in their reactivity; this is probably also true for deep-sea manganese nodules, which have a poorly crystallized structure.

This reaction between nodules and aqueous formaldehyde oxime seems promising, as a rapid and more or less selective extraction of the valuable elements is attainable, even at low temperature. So far, however, only the leaching operation has been studied in our laboratory, and attention will be given to the recovery of the dissolved metals and the recycling of the oxime.

I thank Professor George Panou, Director of the Department, for fruitful discussions.

Received 1 August; accepted 28 October 1986.

1. Bartušek, M. & Okač, A. *Colln Czech. chem. Commun.* **26,** 52–58, 883–887, 2174–2188 (1961).
2. Marczenko, Z. *Bull. Soc. chim. Fr.* **5eS,** 939–944 (1964).
3. Haynes, B. W. *et al. Bull. Dep. Int. Bur. Mines* no. 679 (1985).
4. Bečka, J. & Jokl, V. *Colln Czech. chem. Commun.* **36,** 3263–3274 (1971).
5. Kirk, W. S. in *U.S. Bur. Mines Bull.* **675,** 176 (1985).

Homoeobox gene expression in mouse embryos varies with position by the primitive streak stage

Stephen J. Gaunt, J. Ross Miller,
Donald J. Powell & Denis Duboule*

AFRC Institute of Animal Physiology, Department of Molecular Embryology, Babraham, Cambridge CB2 4AT, UK
* Laboratoire de Genetique Moleculaire des Eucaryotes, 11 Rue Humann, Strasbourg, France

Pattern formation in animal development requires that genes be expressed differentially according to position in the sheets of cells that make up the early embryo. The homoeobox-containing genes of *Drosophila* are control genes active both in the establishment of a segmentation pattern[1,2] and in the specification of segment identity[2,3]. *In situ* hybridization experiments[4–6] confirm that these genes are expressed in a segmentally-restricted manner and that their expression presages morphological differentiation of segmental structures. Homoeobox genes have recently been isolated from the mouse[7–12] and have been shown to be expressed during mouse development[9–13]. Using *in situ* hybridization, we show here that expression of the mouse homoeobox gene *Mo-10* (ref. 7) is spatially restricted in the developing embryo and that localization of expression is already evident within the germ layers before their morphological differentiation. These findings support the suggestion[14] that the homoeobox genes of mammals, like those of *Drosophila*, may be important in pattern formation.

To examine expression of the *Mo-10* gene (also known as *Hox*-1, ref 9), the 660-bp *Hin*dIII-*Eco*RI fragment containing the homoeobox was subcloned into pSP65 and used to produce high specific activity radioactive RNA probes complementary to *Mo-10* messenger RNA (Fig. 1). Two separate probes were made. Probe 1 contains 144 nucleotides from the homoeobox and ~70 nucleotides of additional 3′ sequence. Probe 2 resembles probe 1, but includes the whole of the homoeobox plus ~400 nucleotides of 5′ sequence. Both probes gave similar results in the *in situ* hybridization experiments.

The $9\frac{1}{2}$ day-old embryo was the most advanced stage to be examined. By this age most of the organ systems of the body are already laid down (Fig. 2*a*). The most intense labelling is seen in the nervous system (Fig. 2*b*). Anterior parts of the neural tube, developing at this stage as the forebrain and midbrain regions are not labelled above background, but intense labelling begins posterior to a point in the hindbrain region just anterior to the otic vesicle, at the level of the second (hyoid) visceral arch. The whole length of the neural tube posterior to this point is labelled, but the concentration of grains declines in more posterior regions. We examined the boundary between labelled and unlabelled parts by phase contrast optics, but found no corresponding morphological change in the neural tissue. At this stage of development metameric segmentation of the hindbrain is apparent in its serial constrictions to form neuromeres. The boundary between labelled and unlabelled neural tissue coincides with the boundary between two neuromeres.

The pattern of labelling observed in non-neural tissue of the $9\frac{1}{2}$ d embryo was less remarkable. Generally, labelling was above background in most posterior mesodermal structures, such as the mesonephric kidney and the somite blocks (not shown), but the labelling was weak compared to that of the nervous system. Mesoderm-derived tissues in the head (Fig. 2*b*), the heart, the liver, the gut, and the appendages all showed only background labelling.

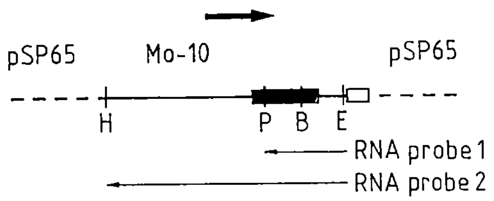


Fig. 1 Use of a pSP65 plasmid[15] containing the *Mo*-10 homoeobox to prepare radioactive RNA probes for *in situ* hybridization and Northern blotting. H, *Hin*dIII, P, *Pvu*II, B, *Bgl*II, E, *Eco*RI. The homoeobox and the SP6 promotor are shown as filled and open boxes respectively. The normal direction of transcription of the *Mo*-10 DNA is from left to right (bold arrow) based upon the conserved open reading frame[7] and homology with the known direction of transcription of other homoeobox genes[10–12].

Methods. To construct the plasmid, the *Hin*dIII-*Eco*RI fragment containing the homoeobox was obtained from a cosmid clone of mouse genomic DNA[16]. The nucleotide sequence (not shown) of this homoeobox is more then 99% homologous to the published sequence[7]. Radioactive RNA probes complementary to putative *Mo*-10 mRNA were prepared after digestion of plasmid DNA with either *Pvu*II (probe 1) or *Hin*dIII (probe 2), followed by *in vitro* transcription using SPS6 RNA polymerase (Promega Biotec). Labelled ribonucleotide triphosphates used were: for *in situ* hybridization, ^{3}H-UTP and ^{3}H-CTP (Amersham, 20 μCi of each, carrier-free, per 20 μl reaction), for Northern blotting, ^{32}P-UTP (Amersham, 50 μCi, 0.2 μg per 20 μl reaction). Final specific activities were about 9×10^7 d.p.m. per μg for ^{3}H-labelled probes and 2×10^8 d.p.m. per μg for ^{32}P-labelled probes. Probes for *in situ* hybridization were subjected to limited alkaline hydrolysis as described[17] for either 45 min (probe 1) or 90 min (probe 2). Probes were then checked by gel electrophoresis to ensure that the bulk of the RNA was reduced to 50–150 nucleotides. A control probe of opposite sense to the probes described above was prepared after subcloning the *Hin*dIII-*Eco*RI fragment of *Mo*-10 DNA into pGEM-4 (Promega Biotec), then using T7 RNA polymerase in the transcription reaction. This probe, hydrolysed as above, did not produce specific labelling when used in the *in situ* hybridization experiments.

At $7\frac{1}{2}$ d of gestation, the cells that will form the mouse are arranged as the three germ layers: ectoderm, mesoderm and endoderm (Fig. 3*a*). But in the posterior region (the site of the primitive streak) ectoderm and mesoderm layers remain unseparated. By $7\frac{3}{4}$ d (Fig. 3*c*), anterior ectoderm in the region of the head fold has become thickened and elevated, presaging the formation of the brain. *In situ* hybridization experiments on embryos at these stages show that labelling is restricted to posterior regions, including the primitive streak (Fig. 3*b*) and adjacent regions of ectoderm and mesoderm (Fig. 3*d*). The anteroposterior boundary between unlabelled and labelled regions is not sharply defined at these stages of development. From examination of twelve $7\frac{1}{2}$ d to $7\frac{3}{4}$ d embryos cut in sagittal section, it was found that labelling varied from about one-third to two-thirds of the length of the embryo, with older embryos tending to show the greater extent of labelling. The extraembryonic tissues at these stages showed only background labelling.

In Northern blot analysis of poly(A)$^+$ RNA extracted from $9\frac{1}{2}$ d and $10\frac{1}{2}$ d whole mouse embryos, probe 1 hybridized with two distinct transcripts of 4 kilobases (kb) and 1.7 kb (Fig. 4). The bands were detected after washing the filter under highly stringent conditions, suggesting that both transcripts are products of the *Mo-10* gene.

The *in situ* hybridization results are consistent with the hypothesis[14] that the homoeobox genes of mammals, like those of *Drosophila*, may be important in the specification of body pattern. The two main observations which suggest such a role for the *Mo-10* gene are, first, that its expression in embryos is spatially restricted in tissues that appear uniform in cell composition on the basis of morphology alone and second, that spatial restriction in expression is evident very early in development, before any observable signs that the body pattern has already been laid down. Since precise fate maps are not available for mouse embryos, it is not possible to conclude that the labelled

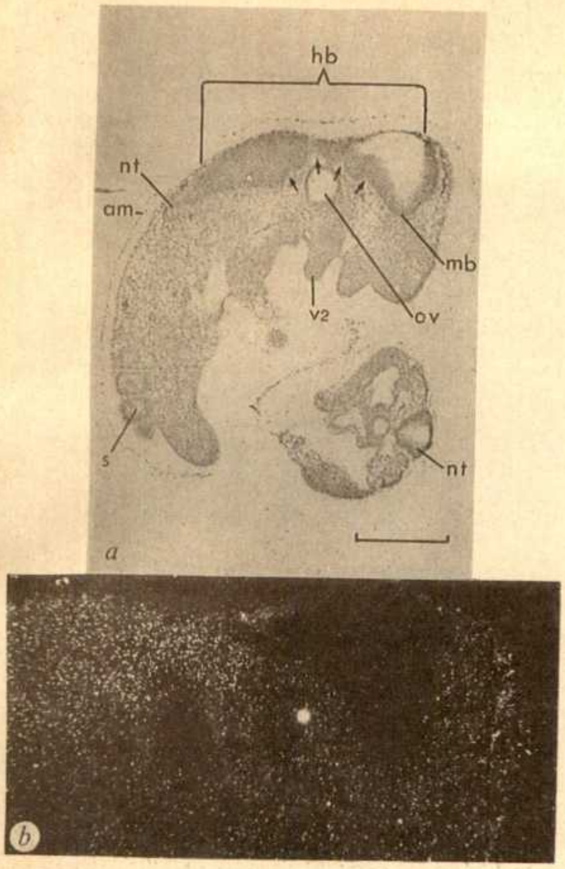


Fig. 2 Localization of *Mo*-10 transcripts, using probe 2 (Fig. 1), by *in situ* hybridization on sections of the $9\frac{1}{2}$ d mouse embryo. *a*, Embryo cut in parasaggital section, under bright-field illumination. hb, hindbrain; mb, midbrain; nt, neural tube; ov, otic vesicle; v2, second visceral arch; s, somite; am, amnion; arrows, constrictions between neuromeres. Scale bar, 0.5 mm. *b*, Silver grains in the head region of the same section, viewed under dark-field illumination. The most intense labelling was detected in the nervous system, but labelling was not seen in anterior regions of the brain. The boundary between labelled and unlabelled neural tissue was located in the hindbrain at the junction of two neuromeres.

Methods. Embryos were obtained from natural matings between F_1 (CBA × C57Bl/6) mice. The morning that the copulation plug was found in the vagina was designated day $\frac{1}{2}$ of pregnancy. After dissection from the uterus into phosphate buffered saline, the embryos were fixed in Bouins fluid for 24 h at 4 °C. They were then dehydrated in ethanol, cleared in xylene, and embedded in paraffin wax (m.p. 56 °C). Sections (10 μm thick) were mounted on poly-L-lysine treated glass microscope slides, dried at 37 °C for 4 h, de-waxed in xylene, rehydrated in decreasing ethanol concentrations and digested with proteinase K as described[17]. Treatment with acetic anhydride and the conditions of hybridization, digestion with RNAase A and washing were as described[18]. After the final wash, sections were immersed in 0.3 M ammonium acetate in 95% ethanol and allowed to dry in air. Slides were dipped at 50 °C in Ilford L4 emulsion diluted 1 : 2 with water, allowed to dry and stored at 4 °C. After an exposure time of three weeks, the autoradiograms were developed in Kodak D19 developer for 3 min at 20 °C. The tissue sections were then stained in Giemsa, dehydrated in ethanol and mounted under coverslips in DPX mountant.

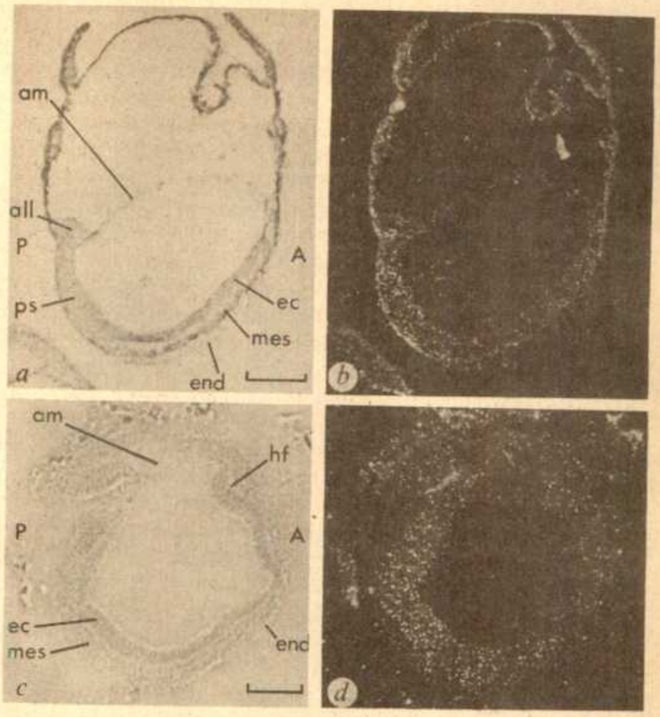


Fig. 3 *In situ* hybridization to detect *Mo*-10 transcripts in sections of primitive streak stage mouse embryos. *a*, *b*, $7\frac{1}{2}$d embryo in sagittal section examined with probe 2; *c*, *d*, $7\frac{3}{4}$ d embryo in parasagittal section, probe 1. *a* and *c* show sections by bright-field and phase contrast illumination respectively. A, anterior; P, posterior; ps, primitive streak; ec, ectoderm; mes, mesoderm; end, endoderm; am, amnion; all, allantois: hf, head fold. Scale bars, 0.1 mm. *b* and *d* show the sections by dark-field illumination. Labelling was restricted to the posterior regions of the embryonic tissue, including the primitive streak (*b*), ectoderm and mesoderm (*d*). The bright border seen at some positions along the edge of the embryonic and extraembryonic tissues in *b* is not due to silver grains, but to a light scattering effect by material which was often present at this stage. Methods were as for Fig. 2.

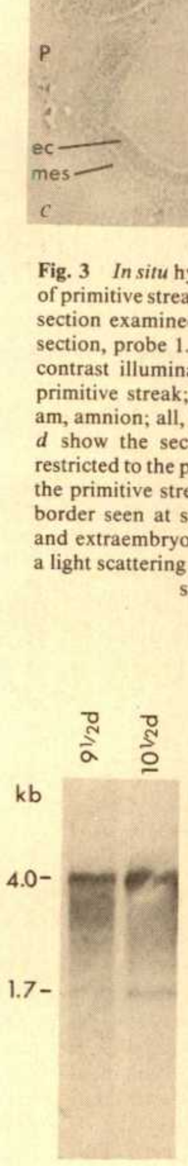


Fig. 4 Northern blot hybridization of *Mo*-10 probe 1 (Fig. 1) to poly(A)$^+$ RNA isolated from whole mouse embryos of $9\frac{1}{2}$ and $10\frac{1}{2}$ d gestation.

Methods. RNA was prepared as described[19]. About 5 μg of each poly (A)$^+$ RNA sample was separated by electrophoresis through a formaldehyde-agarose gel. RNA was then blotted on to Gene Screen Plus. The filter was prehybridized for 2 h at 60 °C in 1 M NaCl, 50 mM Tris HCl *p*H 8.0, 1% SDS, with 100 μg ml^{-1} *Escherichia coli* transfer RNA and 10 μg per ml sonicated salmon sperm DNA. Hybridization with ^{32}P-labelled probe 1 was carried out overnight under the same conditions but with 10% dextran sulphate added. The concentration of probe was about 2 ng ml^{-1}. The filter was washed twice in 2 × SSC, 1% SDS at 65 °C for 20 min, then once in 0.1 × SSC, 1% SDS at 65 °C for 20 min before autoradiography.

ectoderm tissue seen in the $7\frac{1}{2}$ d embryo would normally have given rise exactly to the labelled neural tissue seen at $9\frac{1}{2}$ d. Before drawing such a conclusion it would be necessary to examine several intermediate stages since the primitive streak region is one of considerable cell proliferation and migration.

It is possible that the spatial restriction of gene expression that we have described is not primarily concerned with pattern formation, but instead reflects some other developmental difference between the trunk and head regions of the mouse. However, some observations recently described for the mouse homoeobox gene *Mo-EA*[1,12] suggest further parallels between patterns of expression of homoeobox genes in *Drosophila* and in mice. Awgulewitsch *et al.*[12] used *in situ* hybridization to localize *Mo-EA* transcripts in the nervous system of newborn mice. Labelling was absent in anterior regions of the nervous system but began at about mid-cervical region. Labelling was most concentrated at this point, becoming less intense in more posterior regions of the spinal cord. Thus, *Mo-EA* transcripts, like *Mo-10* transcripts, are detected only posterior to a discrete boundary in the nervous system. This boundary is in the mid-cervical region for *Mo-EA* and in the hindbrain region for *Mo-10*. It should be remembered that transcription of these two genes was examined at different stages in the development of the mouse, but the apparent distinction in the sites of transcription between the two genes is reminiscent of the localized expression of *Drosophila* homoeobox genes within the ventral nerve cord of the developing fly[5,6,20].

We cannot rule out the possibility that the minor 1.7 kb transcript detected in our Northern blots might be due to cross-hybridization of the *Mo-10* probe with a product of another homoeobox gene. But we think it is more likely that the two transcripts result from differential processing of a single gene, as suggested for other mouse homoeobox genes[9,13]. Variability in transcript size could then result from use of alternative initiation, termination, or splice sites. The transcription patterns of the *Mo-10* gene will become amenable to more detailed analysis when complementary DNA probes become available. We consider it unlikely that our *in situ* hybridization experiments could detect any transcripts not shown on the Northern blots since the conditions of hybridization used *in situ* were more stringent than those used in Northern blotting[21].

The potential of the *Mo-10* gene for a role in pattern formation is clear. For example, gene expression at an abrupt boundary within the hindbrain might serve as a cue, instructing cells immediately posterior to this point to develop differently from those cells anterior to the boundary. Such a mechanism would be similar to that proposed for homoeobox genes in the bithorax complex of *Drosophila*[3,22]. From our findings, it is apparent that *Mo-10* gene expression could serve as a positional cue in this way in both ectoderm and mesoderm-derived structures.

We thank Paul Sharpe for the gift of embryo RNA. D.D. was supported by an EMBO fellowship.

Received 19 May; accepted 14 October 1986.

1. Kuroiwa, A., Hafen, E. & Gehring, W. J. *Cell* **37,** 825–831 (1984).
2. Wakimoto, B. T. & Kaufman, T. C. *Devl biol.* **81,** 51–64 (1981).
3. Lewis, E. B. *Nature* **276,** 565–570 (1978).
4. Hafen, E., Kuroiwa, A. & Gehring, W. J. *Cell* **37,** 833–841 (1984).
5. Levine, M., Hafen, E., Garber, R. L. & Gehring, W. J. *EMBO J.* **2,** 2037–2046 (1983).
6. Akam, M. E. & Martinez-Arias, A. *EMBO J.* **4,** 1689–1700 (1985).
7. McGinnis, W., Hart, C. P., Gehring, W. J. & Ruddle, F. H. *Cell* **38,** 675–680 (1984).
8. Colberg-Poley, A. M., Voss, S. D., Chowdhury, K. & Gruss, P. *Nature* **314,** 713–718 (1985).
9. Hart, C. P., Awgulewitsch, A., Fainsod, A., McGinnis, W. & Ruddle, F. H. *Cell* **43,** 9–18 (1985).
10. Hauser, C. A. *et al. Cell* **43,** 19–28 (1985).
11. Jackson, I. J., Schofield, P. & Hogan, B. *Nature* **317,** 745–748 (1985).
12. Awgulewitsch, A., Utset, M. F., Hart, C. P., McGinnis, W. & Ruddle, F. H. *Nature* **320,** 328–335 (1986).
13. Colberg-Poley, A. M. *et al. Cell* **43,** 39–45 (1985).
14. McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. *Cell* **37,** 403–408 (1984).
15. Melton, D. A. *et al. Nucleic Acids Res.* **12,** 7035–7056 (1984).
16. Duboule, D., Baron, A., Mahl, P. & Galliot, B. *EMBO J.* **5,** 1973–1980 (1986).
17. Cox, K. H., DeLeon, D. V., Angerer, L. M. & Angerer, R. C. *Devl. Biol.* **101,** 485–502 (1984).
18. Ingham, P. W., Howard, K. R. & Ish-Horowicz, D. *Nature* **318,** 439–445 (1985).
19. Derman, E. *et al. Cell* **23,** 731–739 (1981).
20. Harding, K., Weeden, C., McGinnis, W. & Levine, M. *Science* **229,** 1236–1242 (1985).
21. McConaughy, B. L., Laird, C. D. & McCarthy, B. J. *Biochemistry* **8,** 3289–3295 (1969).
22. Lawrence, P. A. & Morata, G. *Cell* **35,** 595–601 (1983).

Differential and stage-related expression in embryonic tissues of a new human homoeobox gene

Fulvio Mavilio*, Antonio Simeone†, Adele Giampaolo*, Antonio Faiella†, Vincenzo Zappavigna*, Dario Acampora†, Giancarlo Poiana‡, Giovanni Russo§, Cesare Peschle*‖ & Edoardo Boncinelli†

* Department of Hematology, Istituto Superiore di Sanità,
Viale Regina Elena 299, 00161 Rome, Italy
† International Institute of Genetics and Biophysics, CNR,
Viale Marconi 10, 80100 Naples, Italy
‡ Department of Cellular and Developmental Biology, University of Rome, La Sapienza, Piazzale Aldo Moro 1, 00161 Rome, Italy
§ Division of Obstetrics and Gynecology, Ospedale Civile,
83100 Avellino, Italy

The homoeobox is a 183 base-pair (bp) DNA sequence[1] conserved in several *Drosophila* genes controlling segmentation and segment identity[2–4]. Homoeobox sequences have been detected in the genome of species ranging from insects and anellids to vertebrates[5,6] and homoeobox containing genes have been cloned from *Xenopus*[7–9], mouse[10–22] and man[23]. We recently isolated human homoeobox containing complementary DNA clones, that represent transcripts from four different human genes[24]. One clone (HHO.c10) is selectively expressed in a 2.1 kilobase (kb) polyadenylated transcript in the spinal cord of human embryos and fetuses 5–10 weeks after fertilization[25]. We report the characterization of a second cDNA clone, termed HHO.c13, that represents a new homoeobox gene. This clone encodes a protein of 255 amino-acid residues, which includes a pentapeptide, upstream of the homoeo domain, conserved in other *Drosophila*, *Xenopus*, murine and human homoeobox genes. By Northern analysis HHO.c13 detects multiple embryonic transcripts, which are differentially expressed in spinal cord, brain, backbone rudiments, limb buds and heart in 5–9-week-old human embryos and fetuses, in a striking organ- and stage-specific pattern. These observations suggest that in early mammalian development homoeobox genes may exert a wide spectrum of control functions in a variety of organs and body parts, in addition to the spinal cord.

The nucleotide sequence of clone HHO.c13 is shown in Fig. 1*a*. Multiple stop codons are present in each frame up to nucleotide 230. Translation of the longest reading frame from the ATG codon at base 250 would produce a peptide sequence of 255 amino-acid residues, containing a homoeo domain extending up to 41 residues upstream from the carboxy terminus. The putative 285-bp 3′-untranslated (3′-UT) region includes a polyadenylation signal (AUUAAA) 18 nucleotides upstream from the poly(A) tail. As shown in Table 1, the sequence of HHO.c13 homoeobox is very similar (~60–75% nucleotide homology) to that of other homoeoboxes. At the level of protein sequence, the HHO.c13 homoeo domain shows particularly good homology with murine *Hox1-3* (98.3%), *Xenopus Xhox-1A* (93.4%), human HHO.c10 (90.1%) and *Drosophila Deformed* (*Dfd*) (90.2%) domain. Interestingly, the pentapeptide Leu-Pro-Asn-Thr-Lys immediately downstream from the HHO.c13 homoeo domain (Fig. 1*a*) is conserved in the *Dfd*[26,27], *Xhox-1A*[9] and *Hox1-3*[20–22] genes.

The putative protein encoded by HHO.c13 contains the pentapeptide Val-Tyr-Pro-Trp-Met, starting at amino-acid position 133 (−21 from the homoeo domain). As shown in Fig. 1*b*, this sequence is very similar to the Ile-Tyr-Pro-Trp-Met peptide in two other human unrelated cDNA clones containing homoeobox sequences, termed HHO.c1 and c8 (ref. 24). Its presence in HHO.c10 (ref. 25) cannot be assessed, because this clone is truncated within the homoeobox. The pentapeptide in HHO.c13 is conserved in *Xenopus Xhox-1A* gene[9], whereas the pentapeptide in HHO.c1 and c8 is conserved in *Drosophila Dfd*[26], and mouse m6 (P. Gruss and M. Kessel, personal communication) genes. A highly homologous Leu-Tyr-Pro-Trp-Met

---

‖ To whom correspondence should be addressed.

a

```
CATTAATATCTGGCAGGGGCTCTCAAATGTGCCATAGCAAGCTACTTGATTACACGTATGTTATTTAGTTAAATTTGTG      79

AAAATTATGAGATGCTCACCAACCCGGTGATAAACTTGCTCCCTCGCCATTGGCTGGCCTGGTCACATGGCTGCCCAAC     158

                            Apa I
TTTATTCAGTTGACAGCAAGTAGGAGGGCCCTATGGAAGGAGAAAAAAAGACAACACGAGAAAAATTACTATTTTCTAC     237

            →       →               →           10
           met val met ser ser tyr met val asn ser lys tyr val asp pro lys phe
CTTCTGAAATTAATG GTC ATG AGT TCG TAT ATG GTG AAC TCC AAG TAT GTG GAC CCC AAG TTC     300

         20                                     30
pro pro cys glu glu tyr leu gln gly gly tyr leu gly glu gln gly ala asp tyr tyr
CCT CCG TGC GAG GAG TAT TTG CAG GGC GGC TAC CTA GGC GAG CAG GGC GCC GAC TAC TAC     360

         40                                     50
gly gly gly ala gln gly ala asp phe gln pro pro gly leu tyr pro arg pro asp phe
GGC GGC GGC GCG CAG GGC GCA GAC TTC CAG CCC CCG GGG CTC TAC CCA CGG CCC GAC TTC     420

         60                                     70
gly glu gln pro phe gly gly ser gly pro gly pro gly ser ala leu pro ala arg gly
GGT GAG CAG CCT TTC GGA GGC AGC GGC CCC GGG CCT GGC TCG GCG CTG CCT GCG CGG GGT     480

         80                                     90
his gly gln glu pro gly gly pro gly gly his tyr ala ala pro gly glu pro cys pro
CAC GGA CAA GAG CCA GGC GGC CCC GGC GGT CAC TAC GCC GCT CCA GGA GAG CCT TGC CCA     540

        100                                    110
ala pro pro ala pro pro pro ala pro leu pro gly ala arg ala tyr ser gln ser asp
GCT CCC CCG GCG CCT CCG CCG GCG CCC CTG CCT GGC GCC CGG GCC TAC AGT CAG TCC GAC     600

        120                                    130
pro lys gln pro pro ser gly thr ala leu lys gln pro ala val val tyr pro trp met
CCC AAG CAG CCG CCC TCC GGG ACG GCA CTC AAG CAG CCG GCC GTG GTC TAC CCC TGG ATG     660

        140      Eco RI    ▽                   150
lys lys val his ala asn ser val asn pro asn tyr thr gly gly glu pro lys arg ser
AAG AAG GTG CAC GCG AAT TCG GTG AAC CCC AAC TAC ACC GGT GGG GAA CCC AAG CGG TCC     720

        160                                    170
arg thr ala tyr thr arg gln gln val leu glu leu glu lys glu phe his phe asn arg
CGA ACG GCC TAC ACC CGG CAG CAA GTC CTA GAA CTG GAA AAA GAA TTT CAT TTT AAC AGG     780

        180                                    190
tyr leu thr arg arg arg arg ile glu ile ala his thr leu cys leu ser glu arg gln
TAT CTG ACA AGG CGC CGT CGG ATT GAA ATC GCT CAC ACC CTG TGT CTG TCG GAG CGC CAG     840

     Bgl II                                    210                  *   *   *
ile lys ile trp phe gln asn arg arg met lys trp lys lys asp his lys leu pro asn
ATC AAG ATC TGG TTC CAG AAC CGG AGG ATG AAG TGG AAA AAA GAT CAT AAG CTG CCC AAC     900

 *   *  220                                    230
thr lys gly arg ser ser ser ser ser ser ser ser ser cys ser ser ser val ala pro
ACT AAA GGC AGG TCA TCG TCC TCA TCT TCC TCC TCA TCT TGC TCC TCC TCA GTC GCC CCC     960

        240                                    250
ser gln his leu gln pro met ala lys asp his his thr asp leu thr thr leu trm
AGC CAG CAT TTA CAG CCG ATG GCC AAA GAC CAC CAC ACG GAC CTG ACG ACC TTA TAGAAGA    1021

         Apa I
GGGGACCCTGGGCCCATCTCTCCCTGCGCACCAGGCTGAGCCGAAGCTGCGGGGCAGGCCGGGCCTGCTGTCACCTCGC    1100

TGGGCTCTAAGGTACTCCGGGGTGGACCTGGGACAAGCAGGCCGCCCTCGGACTAGGTTAGCATCCTGCCCGAGGCAGT    1179

CCCCCTCCCTAGAGCGGGATGGGGATGGGAGGGGGGGCGGGATTCTCTCTCTAAGTATGTTATATGGCAGGAGCTACTG    1258

AGAACATAAAATCTTGGCGAGTCATTAAACTTATGAAAAGG                                          1299
```

b

```
HHO.c13         Leu Lys Gln Pro Ala Val Val TYR PRO TRP MET Lys Lys Val His Ala Asn .. 10 AA .. HOMOEO DOMAIN
HHO.c1 (24)     Ala Glu Ser Asn Phe Arg Ile TYR PRO TRP MET Arg Ser Ser Gly Thr Asp .......... HOMOEO DOMAIN
HHO.c8 (24)     Gln Lys Ala Ser Ile Gln Ile TYR PRO TRP MET Gln Arg Met Asn Ser His ... 8 AA .. HOMOEO DOMAIN
Xhox-1A (9)     Pro Glu Gln Asp Pro Val Val TYR PRO TRP MET Lys Lys Ala His Ile Ser .. 10 AA .. HOMOEO DOMAIN
Antp (28)       Ser Gly Met Pro Ser Pro Leu TYR PRO TRP MET Arg Ser Gln Phe Gly Lys ... 3 AA .. HOMOEO DOMAIN
Dfd (27,26)     Thr Asp Gly Glu Arg Ile Ile TYR PRO TRP MET Lys Lys Ile His Val Ala .. 13 AA .. HOMOEO DOMAIN
cad (29)        Ser Pro Ser Lys Pro Pro Tyr Phe Asp TRP MET Lys Lys Pro Ala Tyr Pro .. 10 AA .. HOMOEO DOMAIN
```

c

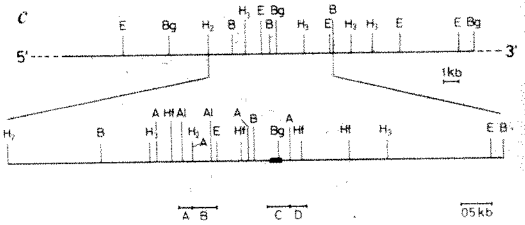


Fig. 1 *a*, Nucleotide sequence of the human cDNA clone *HHO.c13*, with the theoretical translation of its longest open reading frame. Relevant restriction endonuclease recognition sites are also shown. Arrows, three methionine codons, the putative alternative translation start sites: none of them has very good homology with the consensus sequence of initiation sites in eukaryotic genes[37]. The boxed N-terminal amino-acid sequence shows a high degree of homology with the corresponding peptide sequence of *Xenopus Xhox-1A*: Met-Arg-Met-Ser-Ser-Phe-Leu-Ile-Ser-Ser-Asn-Tyr-Val-Asp-Pro-Lys-Phe-Pro-Pro-Cys-Glu-Glu-Tyr (see ref. 9). The sequence encoding the homoeo domain is boxed. The pentapeptide Val-Tyr-Pro-Trp-Met, homologous to that found in a corresponding region of a variety of other homoeobox genes (see Fig. 1*b*) and the polyadenylation signal (ATTAAA)[38] preceding the poly(A) tail are underlined. Asterisks, the conserved pentapeptide downstream from the homoeo domain. ∇, position of the 541-bp intron. *b*, The second small homoeo domain in seven homoeobox genes from man (HHO.c13, c1 and c8), *Drosophila* (*Antp*, *Dfd* and *cad*) and *Xenopus* (*Xhox-1A*). The pentapeptide domain is boxed. Identical amino acids are in capital letters. *c*, Top: restriction endonuclease map of the *HHO.c13* transcription unit on a genomic clone (λ13G), as determined by Southern blot analysis. Bottom: A, B, C, D indicate contiguous subclones of HHO.c13 (*Pst*I/*Apa*I, *Apa*I/*Eco*RI, *Eco*RI/*Apa*I, *Apa*I/*Bam*HI), used as probes; black box, homoeobox. A, *Apa*I; Al, *Alu*I; B, *Bam*HI; Bg, *Bgl*II; E, *Eco*R1; H2, *Hin*dII; H3, *Hin*dIII; Hf, *Hin*fII; kb, kilobase pairs.

Methods. Human cDNA clones containing homoeobox sequences from simian virus 40 (SV40)-transformed human fibroblasts[24] were subcloned in the pEMBL8 plasmid[39] and sequenced as described by Maxam and Gilbert[40]. A genomic DNA clone (λ13G) was isolated from a human genomic library in λ Charon 28, digested with restriction endonucleases, electrophoresed on 0.8–1.5% agarose gels, transferred to nitrocellulose paper by Southern blotting and hybridized to nick-translated HHO.c13 clone or subclones according to standard techniques[41].

sequence is present in the *Drosophila Antp* gene[28]. The pentapeptides are located upstream from the homoeo domains starting between positions −11 and −24 (Fig. 1*b*). Similarly, a Trp-Met dipeptide is found at −18 from the unique homoeo domain of *Drosophila caudal* (*cad*) gene[29]. The pentapeptide apparently defines a second small homoeo domain, located at a short, slightly variable distance from the 5′ end of the homoeo domain itself, that is not encoded by the homoeobox containing exon[9,26,28,29].

Finally, it is noteworthy that the 23 amino-terminal residues deduced from HHO.c13 bear a striking homology to the corresponding peptide sequence encoded by *Xhox-1A*[9] (Fig. 1*a* legend).

HHO.c13 was used as hybridization probe to screen a human genomic library in λ Charon 28. A 24-kb genomic clone (λ13G) containing HHO.c13 was characterized by restriction endonuclease mapping (Fig. 1*c*) and partially sequenced (not shown). The HHO.c13 coding region is interrupted by a 541-bp intron, located upstream from the homoeobox (Fig. 1*a*, *c*).

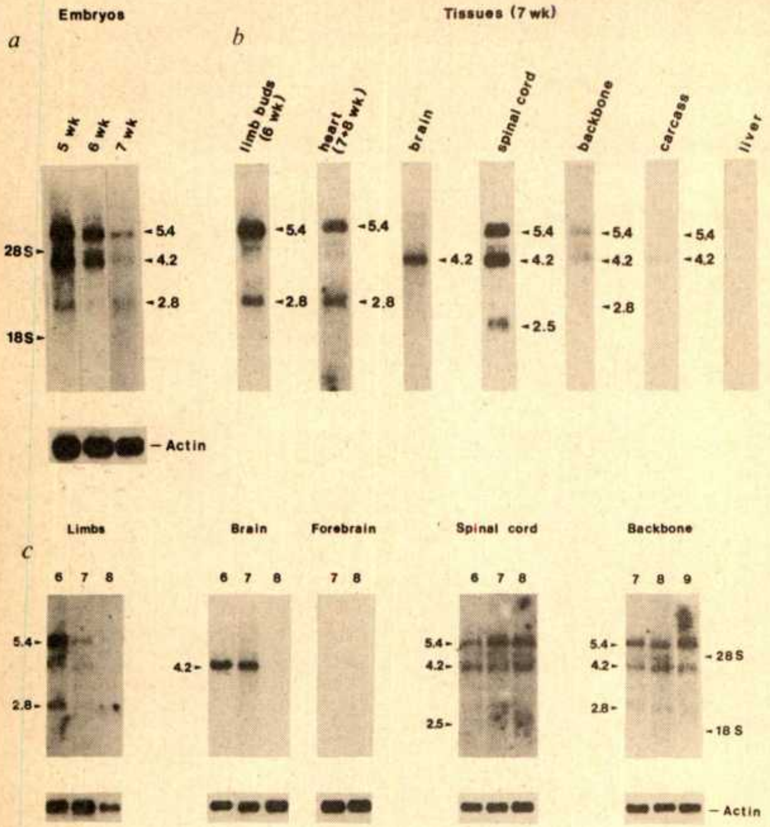


Fig. 2 *a*, Northern blot analysis of the expression of *HHO.c13* in poly(A)$^+$ RNA from whole human embryos at 5, 6 and 7 weeks after fertilization. The relative mobility of 28S and 18S rRNA is shown as size markers. Sizes are in kb. *b*, Expression of HHO.c13 in poly(A)$^+$ RNA obtained from pools of organs or body parts dissected from human embryos of 7 weeks. Limb buds were obtained at 6 weeks, hearts pooled at 7 and 8 weeks. Brain was dissected from spinal cord at the level of the pontine flexure, separating metencephalon from myelencephalon. 'Carcass' indicates poly(A)$^+$ RNA from a control embryo deprived of limbs, brain and spinal column. *c*, Expression of HHO.c13 in limb buds or limbs, brain, spinal cord and backbone from human embryos of 6, 7, 8 and 9 weeks, as analysed by Northern blot hybridization of poly(A)$^+$ RNA. Sizes are in kb. The forebrain contains the cerebral hemispheres, dissected from midbrain (mesencephalon) and hindbrain (metencephalon).

Methods. Total RNA was extracted by the guanidinium thiocyanate technique[42], poly(A)-selected by one passage on oligo (dT)-cellulose columns, run in 2-3 μg aliquots on 1.0% agarose-formaldehyde gels[41] and transferred to nitrocellulose or nylon membranes by Northern capillary blotting[43]. Filters were hybridized to 10^7 c.p.m. of HHO.c13 probe labelled by nick-translation to specific activity of $3\text{-}6\times10^8$ d.p.m. per μg, washed at high stringency (final wash 0.1 × SSC, 0.1% SDS at 60 °C) and exposed for 1-7 days to Kodak XR-5 X-ray films in X-omatic intensifying screen cassette. Filters were rehybridized to a chicken β-actin probe[44] for normalization.

Chromosome mapping by standard procedures indicates that the HHO.c13 gene is localized on the long arm of chromosome 2 (K. Hübner and C. Croce, personal communication). These data indicate that the HHO.c13 clone defines a new human homoeobox gene, unrelated to the cluster previously identified on chromosome 17[12,13].

We then used the HHO.c13 cDNA clone as a probe to analyse the expression of the gene in human embryonic and early fetal development. Embryos and fetuses were obtained virtually intact from legal curettage abortions, from 5-9 weeks after fertilization[30]. The age of embryos (5-8-week-old) was established by morphological staging according to multiple criteria[31]; the youngest were from the first part of the fifth week (total length 7-9 mm, developmental stage 13-14). Fetuses (9-week-old) were staged on the basis of standard age/crown-rump length (CRL) plots[32]. The dating error was about ±2 days (see also refs 25, 30, 31).

Northern blot analysis of poly(A)$^+$ RNA obtained from whole embryos at 5, 6 and 7 weeks (Fig. 2*a*) indicates that the HHO.c13 gene is expressed in at least three transcripts of 5.4, 4.2 and 2.8 kb: their abundance, however, sharply declines between the fifth and seventh weeks. No signal was detected in Northern blots of corresponding poly(A)$^-$ RNA (not shown). Poly(A)$^+$ RNA was probed with cDNA subclones with or without the homoeobox (probes C or B and D respectively, see Fig. 1*c*): in all cases the pattern was virtually identical (not shown), thus excluding cross-hybridization with transcripts from other homoeobox containing genes. Note that the noncoding strand of the central 170-bp *Eco*RI-*Bgl*II fragment of HHO.c13 (see Fig. 1*a*) detects all three major transcripts, whereas the coding strand does not show any major hybridization band (data not shown).

We then analysed the expression of HHO.c13 in different organs or body parts, dissected by microsurgery from 7-week-old embryos (CRL 16-20 mm, stage 18-19); samples from different specimens were pooled when necessary. Limb buds were obtained at 6 weeks (CRL 10-15 mm, stage 16-17), hearts pooled at 7-8 weeks (CRL 20-20 mm, stage 20-23). Brain and spinal cord, dissected free of contaminating tissues, were separated at the level of the pontine flexure, which delimits metencephalon from myelencephalon (*medulla oblongata*). As shown in Fig. 2*a*, *b*, human embryos show a differential, tissue-specific pattern of expression of the HHO.c13 gene, both quantitatively and qualitatively. In particular, (1) 5.4- and 2.8-kb transcripts are predominantly expressed in limb buds and heart; (2) only the 4.2-kb band is observed in brain, but (3) 5.4-, 4.2- and 2.5-kb transcripts are detectable in spinal cord, the last showing the characteristic broad appearance of other cord-specific homoeobox-containing transcripts (ref. 25 and our unpublished results); (4) 5.4-, 4.2- and a faint 2.8-kb bands are observed in the backbone rudiment; (5) 4.2- and to a lesser extent 5.4-kb transcripts are detectable in RNA from control embryos ('carcass'), deprived of limbs, brain and spinal column; (6) no signal is observed in liver.

To define whether the expression of HHO.c13 transcripts is stage-specific, we have analysed poly(A)$^+$ RNA from limb buds or limbs, brain, spinal cord and backbone at 6, 7 and 8 weeks (Fig. 2*c*).

Expression of the 5.4- and 2.8-kb transcripts in limbs declines dramatically between the sixth and the seventh weeks, and almost disappears by eight weeks. This correlates with the morphogenetic progression of limb development, which starts late in the fifth week and is almost complete at the end of the eight week[31]. Significant expression of HHO.c13 transcripts is thus confined to the limb bud development stage.

The 4.2-kb transcript observed in the brain is expressed at constant or possibly slightly declining level between six and seven weeks, but almost disappears by eight weeks. Dissection of forebrain (developing cerebral hemispheres) from midbrain (mesencephalon) and hindbrain (methencephalon) indicates that at seven weeks expression of this transcript is confined to the two latter structures. We could not discover whether this spatial restriction also occurs at 6 weeks; if so, it might be related to development of well-defined structures in the 6-7 week period (for example, the cerebellum).

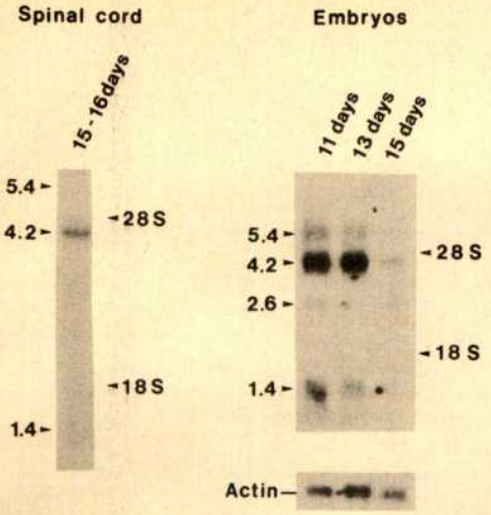


Fig. 3 Northern blot analysis of the expression of a HHO.c13 rat homologue. Poly(A)$^+$ RNA was obtained from both spinal cord at 15–16 days from fertilization (5 μg per lane, left panel) and embryos *in toto* at 11, 13 and 15 days (3 μg, right panel). Relevant transcripts are indicated. Sizes are in kb.

Table 1 Homologies between the HHO.c13 homoeobox and the corresponding region of other genes or clones, at nucleotide and amino-acid level

Genes or clones	Nucleotides (%)	Amino acid (%)	Reference
Human			
HHO.c1	72.0	82.0	24
HHO.c8	72.0	77.0	24
HHO.c10 (=*Hu1*)	74.3	90.1	25, 23
Hu2	68.3	78.7	23
Mouse			
m5	69.9	80.3	18
m6	74.3	82.0	18
Mo10	71.0	75.4	10
Hox 1-3	81.9	98.3	21
Hox 3	66.1	63.9	19
Xenopus			
AC1	68.3	77.0	8
MM3	72.6	83.6	8
Xhox-1A	76.5	93.4	9
Xhox-1B	73.8	90.0	9
Drosophila			
Antp	72.0	83.6	45
ftz	68.3	77.0	45
Ubx	65.6	70.5	45
Dfd	73.2	90.2	27, 26
Scr	72.7	87.9	46, 27, 26
iab-7	58.2	54.1	27
cad	57.9	55.7	29
Sea urchin			
HB1	66.7	77.0	47

Expression of 5.4-, 4.2- and 2.5-kb transcripts in the spinal cord, and 5.4-, 4.2- and 2.8-kb transcripts in the developing backbone, appear constant throughout the 6–8 and 7–9 week period respectively. It is worth noting that expression of at least one other homoeobox gene (HHO.c10 or *Hu-1*) is apparently constant in the spinal cord in the same period of human embryogenesis[25].

Earlier stages of development were explored in the rat. We first analysed poly(A)$^+$ RNA from spinal cord of rat embryos at 15–16 days from fertilization, corresponding to ∼7 weeks in human development. As shown in Fig. 3, HHO.c13 hybridizes to at least three transcripts of 5.4-, 4.2- and 1.4-kb, presumably from a rat gene homologous to HHO.c13. We then extended the analysis to poly(A)$^+$ RNA from whole rat embryos at developmental stages from day 11 (total length 1.5–2.0 mm, 13 somites) to 15 (CRL 9 mm, >40 somites)[33], (approximately the beginning of the fourth and the end of the sixth week of human embryogenesis)[31]. Transcripts of 5.4-, 4.2-, 2.6- and 1.4-kb were observed (Fig. 3), with constant expression from day 11–13 decreasing by 50% at day 15. Thus the HHO.c13 gene is already active in rat (and possibly human) embryos at a stage of development immediately after the complete fusion of the neural tube.

The transcripts from human embryos are significantly longer than the cDNA clone (isolated from a simian virus 40 (SV40)-transformed adult fibroblast cDNA library[34]). Preliminary mapping data on embryonic RNAs (not shown) indicate that the major 5.4- and 4.2-kb transcripts terminate ∼2 kb downstream from the HHO.c13 polyadenylation site, whereas 2.8-kb transcripts do not extend further downstream. If our cDNA clone is the 3′ portion of the 2.8 kb transcript, it may be suggested that this transcript has an unusually long 5′-UT sequence.

In *Drosophila* the homoeobox genes, mainly organized in multigene clusters, are involved in the control of crucial developmental processes, as the determination of the number and fate of individual body segments[1]. In amphibians and mammals homoeobox genes, often organized in multigene clusters[15,18,21,23], might act similarly. Studies on early mammalian development show that a variety of homoeobox genes are expressed in murine[14–18,21,22] and human[25] embryos, often with a time-related pattern[14–18,21,22]. Analysis of single organs from mammalian embryos or newborns reveals selective expression of homoeobox genes in the spinal cord[14,19,25], which typically constitutes a metameric structure: in particular, *Hox-3* is expressed in a spatially restricted pattern[19], reminiscent of that observed for other homoeobox genes in *Drosophila*[35,36].

Our observations suggest that homoeobox genes may exert a wide spectrum of control functions in the embryonic and early fetal development of mammals, not only in the spinal cord, but also in a variety of other organs and body parts that follow basically different patterns of morphogenesis.

We thank Drs G. Cossu and G. Macino for helpful discussion, Dr V. Fidanza for cooperation in computer analysis and Mrs V. Autuori for expert assistance in rat embryo dissection. This work was supported by grants from Consiglio Nazionale delle Ricerche, Progetti Finalizzati 'Ingegneria Genetica e Basi Molecolari delle Malattie Ereditarie' to E.B. and C.P. (grant no. 85.01425.51), 'Oncologia' to E.B., F.M. (85.02244.44) and C.P. (85.02560.44), 'Medicina Preventiva' to C.P. (85.00707.56).

Received 30 May; accepted 14 October 1986.

1. Gehring, W. J. *Cell* **40**, 3–5 (1985).
2. Ouweneel, W. H. *Adv. Genet.* **18**, 179–248 (1976).
3. Garcia-Bellido, A. *Am. Zool.* **17**, 613–629 (1977).
4. Nuesslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795–801 (1980).
5. McGinnis, W. *Cold Spring Harb. Symp. quant. Biol.* **50**, 263–270 (1985).
6. Holland, P. W. H. & Hogan, B. L. M. *Nature* **321**, 251–253 (1986).
7. Carrasco, A. E., McGinnis, W., Gehring, W. J. & De Robertis, E. M. *Cell* **37**, 409–444 (1984).
8. Mueller, M. M., Carrasco, A. E. & De Robertis, E. M. *Cell* **39**, 157–162 (1984).
9. Harvey, R. P., Tabin, C. J. & Melton, D. A. *EMBO J.* **5**, 1237–1244 (1986).
10. McGinnis, W., Hart, C. P., Gehring, W. J. & Ruddle, F. H. *Cell* **38**, 675–680 (1984).
11. Colberg-Poley, A. M., Voss, S. D., Chowdhury, K. & Gruss, P. *Nature* **314**, 713–718 (1985).
12. Rabin, M. *et al. Nature* **314**, 175–178 (1985).
13. Joyner, A. L. *et al. Nature* **314**, 173–175 (1985).
14. Jackson, I. J., Schofield, P. & Hogan, B. *Nature* **317**, 745–748 (1985).
15. Hart, C. P. *et al. Cell* **43**, 9–18 (1985).
16. Hauser, C. A. *et al. Cell* **43**, 19–28 (1985).
17. Joyner, A. L. *et al. Cell* **43**, 29–37 (1985).
18. Colberg-Poley, A. M. *et al. Cell* **43**, 39–45 (1985).
19. Awgulewitsch, A. *et al. Nature* **320**, 328–335 (1985).
20. Wolgemuth, D. J. *et al. EMBO J.* **5**, 1229–1235 (1986).
21. Duboule, D., Baron, A., Maehl, P. & Galliot, B. *EMBO J.* **5**, 1973–1980 (1986).
22. Rubin, M. R. *et al. Science* **233**, 663–667 (1986).
23. Levine, M., Rubin, G. M. & Tjian, R. *Cell* **38**, 667–673 (1984).
24. Boncinelli, E. *et al. Cold Spring Harb. Symp. quant. Biol.* **50**, 301–306 (1985).
25. Simeone, A. *et al. Nature* **320**, 763–765 (1986).
26. Laughon, A. *et al. Cold Spring Harb. Symp. quant. Biol.* **50**, 253–262 (1985).
27. Regulski, M. *et al. Cell* **43**, 71–80 (1985).
28. Schneuwly, S., Kuroiwa, A., Baumgartner, P. & Gehring, W. J. *EMBO J.* **5**, 733–739 (1986).
29. Mlodzik, M., Fjose, A. & Gehring, W. J. *EMBO J.* **4**, 2961–2969 (1985).

30. Peschle, C. *et al. Nature* **313,** 235–238 (1985).
31. Moore, K. L. *The Developing Human* 3rd edn (Saunders, Philadelphia, 1982).
32. Shepard, T. H., Nelson, T., Dakley, G. P. & Lemiere, R. J. in *Monitoring Birth Defect and Environment. The Problem of Surveillance* (eds Hook, E. B., Janerich, D. T. & Porter, I. H.) 29–43 (Academic, New York, 1971).
33. Nicholas, J. S. in *The Rat in Laboratory Investigation* (eds Farris, E. J. & Griffiths, J. Q.) 51–67 (Hafner, New York and London, 1967).
34. Okayama, H. & Berg, P. *Molec. cell. Biol.* **3,** 280–295 (1983).
35. Levine, M., Hagen, E., Garber, R. L. & Gehring, W. I. *EMBO J.* **2,** 2037–2046 (1983).
36. Harding, K., Wedeen, C., McGinnis, W. & Levine, M. *Science* **229,** 1236–1242 (1985).
37. Kozak, M. *Nucleic Acids Res.* **12,** 857–872 (1984).
38. Birnstiel, M. L., Busslinger, M. & Strub, K. *Cell* **41,** 349–359 (1975).
39. Dente, L., Cesareni, G. & Cortese, R. *Nucleic Acids Res.* **11,** 1645–1655 (1983).
40. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 560–564 (1977).
41. Maniatis, T., Fritsch, E. & Sambrook, J. *Molecular cloning: a laboratory manual* (Cold Spring Harbor Laboratory, New York, 1982).
42. Chirgwin, J. M. *et al. Biochemistry* **18,** 5294–5300 (1979).
43. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77,** 5201–5205 (1980).
44. Cleveland, D. W. *et al. Cell* **20,** 95–105 (1980).
45. McGinnis, W., Garber, A. M., Wirz, J., Kuroiwa, A. & Gehring, W. J. *Cell* **37,** 403–408 (1984).
46. Kuroiwa, A. *EMBO J.* **4,** 3757–3764 (1985).
47. Dolecki, G. J. *et al. EMBO J.* **5,** 925–930 (1986).

Cross-regulatory interactions among the gap genes of *Drosophila*

Herbert Jäckle, Diethard Tautz, Reinhard Schuh, Eveline Seifert & Ruth Lehmann*

Max-Planck-Institut für Entwicklungsbiologie, Abt. Biochemie,
Spemannstr. 35/II, D-7400 Tübingen, FRG
* Max-Planck-Institut für Entwicklungsbiologie, Abt. Genetik,
Spemannstr. 35/III, D-7400 Tübingen, FRG

Segmentation of the *Drosophila* embryo involves expression of several classes of zygotic genes in response to positional information of maternal origin[1]. Zygotic expression of the gap genes is thought to be required for the subdivision of the embryo into several units of adjacent segments; pair-rule genes establish a repeat unit length of two segments and segment polarity genes specify parts of the single segment[2–5]. The positional information provided by the three classes of segmentation genes is interpreted for the establishment of segment identity using the homoeotic genes[6–10]. Molecular probes for several segmentation and homoeotic genes reveal a temporal and spatial pattern of gene expression in the young embryo[11–18]. Changes in the pattern of expression of these genes in mutant embryos provide evidence for a network of interactions among homoeotic and pair-rule genes[19–23] that itself depends on the activity of gap genes[24–27]. We show here that the spatial limit of gap gene expression depends on bilateral interactions among gap genes.

Gap mutant embryos lack a number of adjacent segments which define the domains of the wild-type gene requirement. Based on the deletion pattern (Fig. 1), zygotic *hunchback* (*hb*) activity is required for normal establishment of gnathal and thoracic segments (up to T3), *Krüppel* (*Kr*) for thoracic and anterior abdominal segments (T1–A5), and *knirps* (*kni*) for the abdominal segments A1 to A7 (ref. 2). In a first and naive view, one might expect expression of gap genes in overlapping domains which themselves correspond to the deleted regions in the mutant phenotypes, as shown before for members of the pair-rule class of segmentation genes[13,14,18]. According to a theoretical model by Meinhardt[28], however, the expression of gap genes in response to positional information of maternal origin occurs in adjacent, non-overlapping domains. Each domain should be smaller than the deleted region in the mutant embryo. The limits of the regions where gap genes are expressed (at blastoderm stage) would then provide additional positional cues for the expression of other segmentation genes in a hierarchical order. The model predicts further that these borders

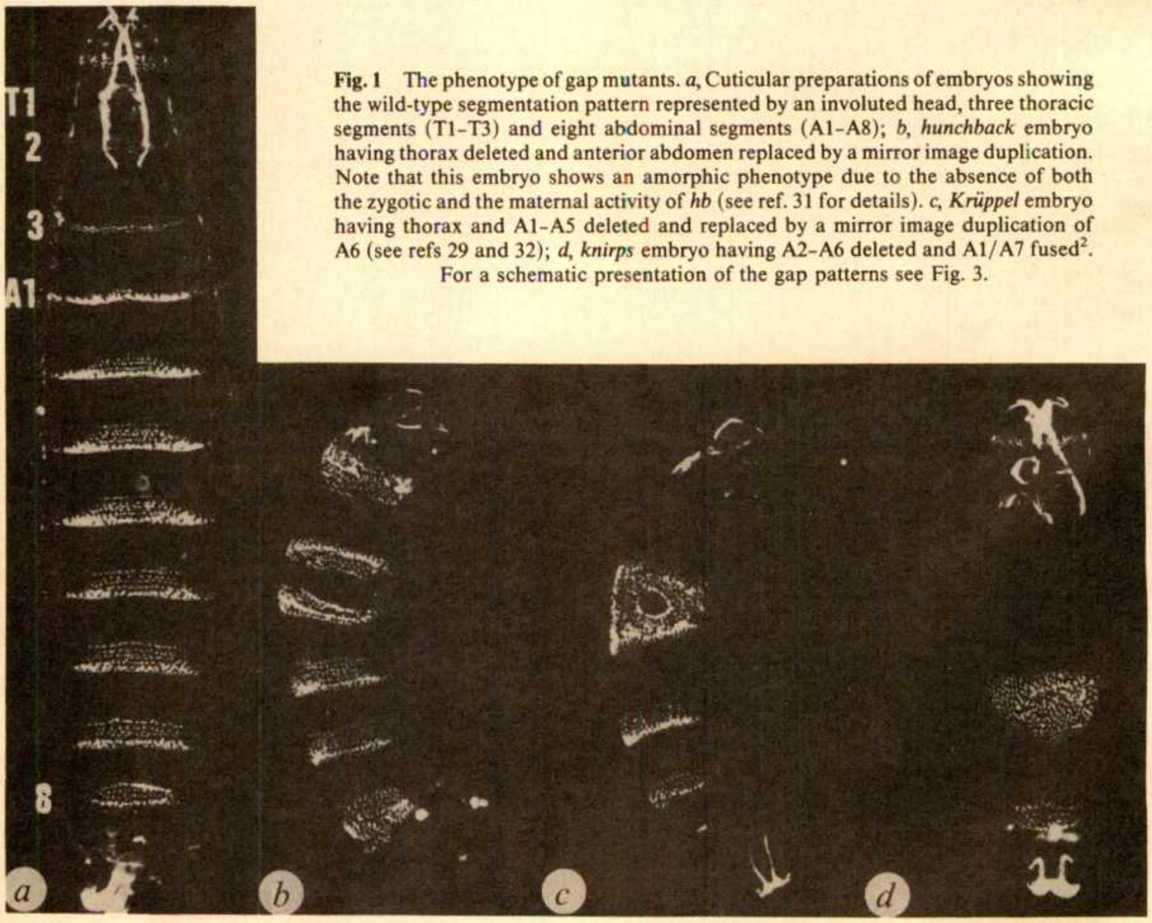


Fig. 1 The phenotype of gap mutants. *a*, Cuticular preparations of embryos showing the wild-type segmentation pattern represented by an involuted head, three thoracic segments (T1–T3) and eight abdominal segments (A1–A8); *b*, *hunchback* embryo having thorax deleted and anterior abdomen replaced by a mirror image duplication. Note that this embryo shows an amorphic phenotype due to the absence of both the zygotic and the maternal activity of *hb* (see ref. 31 for details). *c*, *Krüppel* embryo having thorax and A1–A5 deleted and replaced by a mirror image duplication of A6 (see refs 29 and 32); *d*, *knirps* embryo having A2–A6 deleted and A1/A7 fused[2]. For a schematic presentation of the gap patterns see Fig. 3.

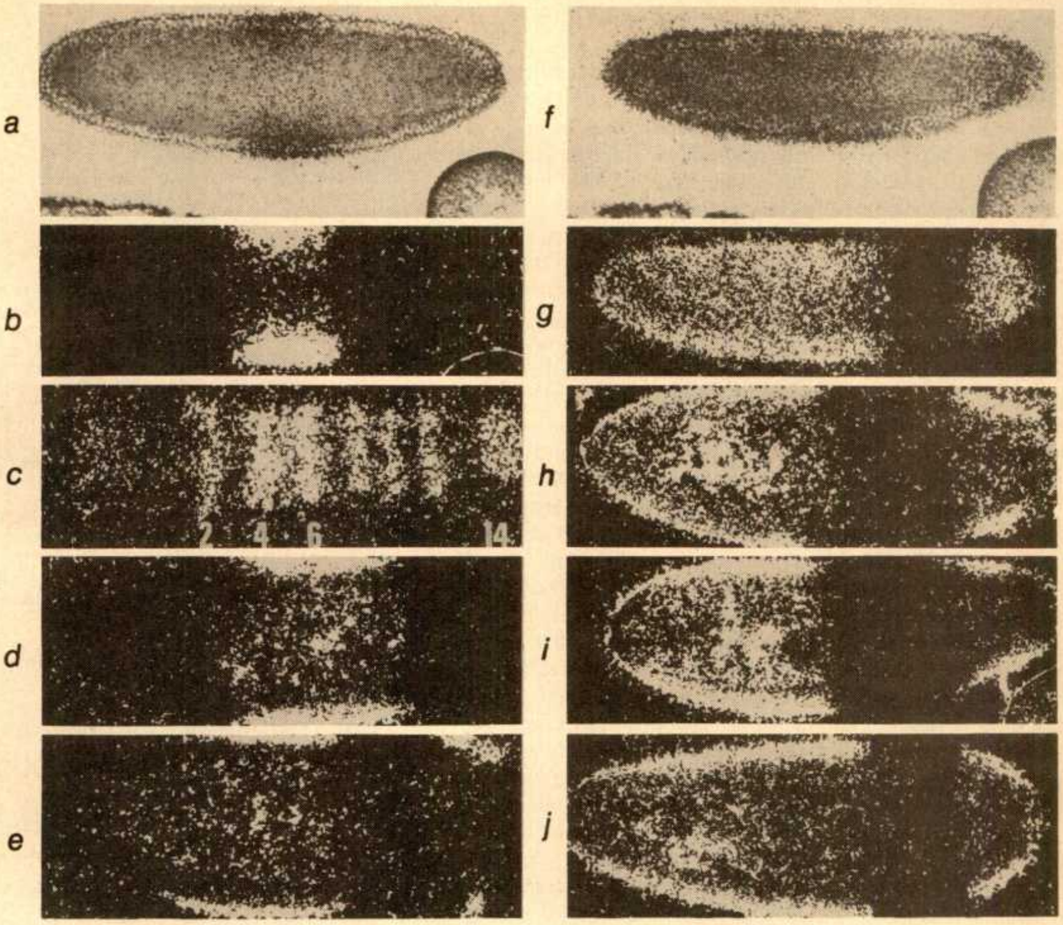


Fig. 2 *In situ* localization of Kr^+ and hb^+ transcripts in gap mutant embryos after the thirteenth nuclear division. *a*, *b*, Kr^+ expression in blastoderm stage wild-type embryos; *a*, bright field photograph of a parasagittal section and *b*, corresponding dark-field photograph to visualize the *in situ* hybridization pattern of the *Kr* probe. Note that even after overexposure, the Kr^+ anterior domain remains within the borders of 12–14 cells[17]. *c*, *In situ* pattern of a mixed *fushi tarazu* (*ftz*) and Kr^+ probe hybridized to the tangential section of the embryo shown in *a*. Note that the Kr^+ domain (low specific activity) overlaps *ftz*-stripes 2 and 4 and thus corresponds to the approximate limits of parasegments 4–6 (indicated by numbers, ref. 14). Kr^+ expression in *d*, kni^- and *e*, hb^- embryos. Note the broadening of Kr^+ expression into the adjacent gap gene domain if the corresponding gene activity is absent. *f*, Bright field and *g*, corresponding dark field photograph of wild type section between *a* and *c* which was hybridized with a mixed hb^+/Kr^+ probe. *h*, parasagittal section (dark field) of the same embryo hybridized with the hb^+ probe. Expression of hb^+ in: *i*, kni^- and *j*, Kr^- embryos. Note a broadening of the hb^+ domain in Kr^- but not in kni^- embryos. Orientation of embryos is anterior left and ventral bottom. Note that yolk cells express both Kr^+ and hb^+ (ref. 17, and D.T., unpublished).

Methods. Embryos from wild-type, *Kr* AP1/SM1 (ref. 29), *kni* FC13/TM3 or *hb* Df(3R)p13/TM3 (ref. 33) parents were collected within 2 h after deposition and kept for 1 h at 26 °C. Serial sections were prepared and subsequently handled as described[17], except that they were hybridized *in situ* using nick-translated ^{35}S-labelled hb^+ and/or Kr^+ probes ($2\text{–}5\times10^7$ c.p.m. per μg DNA). Exposure was 10–20 days in the cold. For positional reference in the embryo, the mixed ftz/Kr^+ probe was hybridized to wild-type and gap mutant sections alternating with Kr^+ and/or hb^+ hybridized sections (see Fig. 3 legend). This allowed, in addition to statistics, hb^- embryos to be recognized by the lack of hb^+ transcripts and a changed ftz^+ pattern[26]; Kr^- by the lack of Kr^+ transcripts and a changed ftz^+ pattern[26]. Embryos which were kni^- were identified by the broadened Kr^+ domain in about a quarter of embryos which have the changed ftz^+ pattern[26]. The limit of resolution of Kr^+ and hb^+ expression on our sections is about two cells.

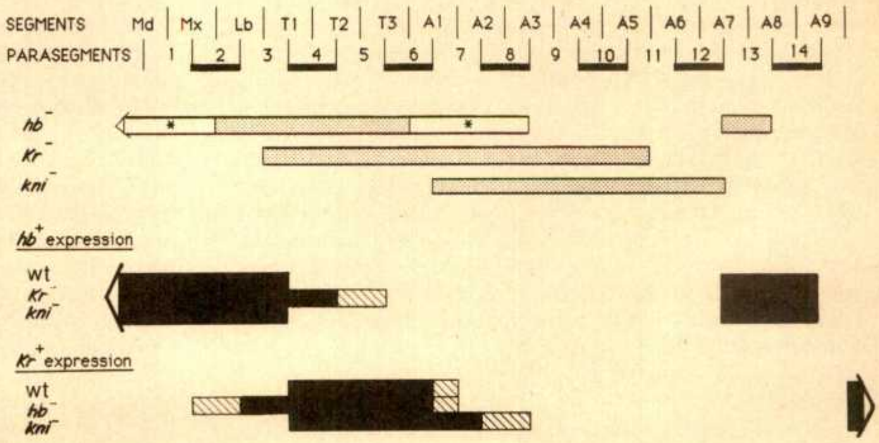


Fig. 3 Summary of the spatial distribution of hb^+ and Kr^+ expression in gap mutant embryos. Schematic diagram of the patterns of hb^+ transcripts in wild type (wt), Kr^- and kni^- embryos, and Kr^+ transcripts in wild type (wt), hb^- and kni^- embryos, respectively. Black bars, areas of expression, dashed areas, subjective limits of resolution. Above, the deletion of adjacent segments in the different gap mutant embryos is indicated by dotted bars. Asterisks, gap extension in *hb* embryos in the absence of the maternal hb^+ complement[31]. For a positional reference of Kr^+ and hb^+ domains in wild-type embryos, segments and parasegments 1–14 were determined by use of the *fushi tarazu* probe, which is localized in parasegments 2, 4, etc. (thin black bars, see Fig. 2*c*) as described[14]. The analysis is within the limits of two cells and it is based on a minimum of 16 embryos with the respective genotype. Md, mandibular segment; Mx, maxillary segment; Lb, labial segment; T1–T3, thoracic segments; A1–A8, abdominal segments.

are established by interaction between pairs of gap genes, the activities of which are maintained in their domains by autocatalytic and competitive processes[28].

Recent cloning of Kr^{29} and *hb* (D.T., unpublished) provides molecular probes to analyse the spatial pattern of Kr^+ and hb^+ expression by *in situ* hybridization to gap mutant embryos (Fig. 2). Adjacent sections of early blastoderm stage embryos were probed with a mixed Kr^+/hb^+ (Fig. 2*f,g*) or Kr^+/*fushi tarazu* probe to provide positional reference points. This allowed us to map the 12-14 cell wide domain of Kr^+ expression in parasegments 4-6 and the anterior domain of hb^+ expression (Fig. 2*j*) anterior to *Kr* (Fig. 2*a–c*; *f-h.*) The size and the borders of the *Kr* and the *hb* domains were changed in gap mutant embryos, but only if the activity of the gap gene in the adjacent domain was missing. In amorphic kni^- embryos, Kr^+ expression extends 6-8 cells posterior to its normal boundary (Fig. 2*d*), into the region corresponding to parasegment 8 or segment A3 in the wild type embryo (Fig. 3). Expression of hb^+, in contrast to *Kr*, appears unchanged in kni^- embryos (Fig. 2*i*). But the hb^+ domain broadens 6-8 cells posteriorly if Kr^+ activity is absent (Fig. 2*j*). It forms a new posterior boundary in a region which is normally parasegment 5 or segment T3 (Fig. 3). Kr^+ expression in hb^- embryos has a complementary effect; the Kr^+ expression extends 6-8 cells anteriorly, if hb^+ activity is missing in the embryo (Fig. 2*e*). The new anterior boundary of Kr^+ appears in a region which corresponds to parasegment 2 or maxilla in normal embryos (Fig. 3). These results indicate that the normal domain of gap gene expression is dependent on gap gene activity in the adjacent domains. This suggests cross-regulation of gap gene activity at the same hierarchical level as has been proposed as a key prediction in Meinhardt's model[28].

There is a second, posterior domain of Kr^+ (at late blastoderm; see details in ref. 17) and hb^+ expression (Fig. 2; and D.T., unpublished) that relates to the fusion of abdominal segments A7 and A8 in late hb^- embryos, and to the lack of malphighian tubules in late Kr^- embryos. The spatial characteristics of these posterior domains remain unaffected in each of the gap mutant embryos analysed (Fig. 2 and data not shown). This suggests that the control of expression in the posterior domain is different from anterior, possibly depending on different maternal cues, other gap genes which may act in that region and/or a feedback control by the activities of other segmentation genes. Such activities may also be required to establish the new anterior and posterior limits of Kr^+ activity in hb^- and kni^- embryos, respectively. Neither of these boundaries have any obvious morphological correspondence in the phenotype of any known gap mutant embryo. Thus the *hb* posterior border in Kr^- embryos may relate to the complementary extension of the kni^+ domain and hence an interaction of *hb* and *kni* in the absence of the Kr^+ gene product, as predicted by Meinhardt[28].

We do not understand the molecular basis of these interactions between gap genes. Because the *Kr* gene product shares homology with the DNA-binding domain of transcription factor[30], it is tempting to speculate that gap gene products may regulate each other's activity through mutual or competitive DNA-binding interactions as has been suggested for cross-regulation of the expression of homoeobox containing genes[21,22].

We thank Drs H. Meinhardt, S. M. Cohen, P. Lawrence, M. Akam, I. Baxivanelis, C. Blume and R. Groemke-Lutz. D.T. was supported by a Deutsche Forschungsgemeinschaft (DFG) fellowship. This work was in part supported by a DFG grant (Leibniz-Program).

Received 26 August; accepted 14 October 1986.

1. Nüsslein-Volhard, C. in *Determinants of Spatial Organization* (eds Subtelney, S. & Konigsberg, I. R.) 185-211 (Academic, New York, 1979).
2. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287,** 795-801 (1980).
3. Nüsslein-Volhard, C., Wieschaus, E. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193,** 267-282 (1984).
4. Wieschaus, E., Nüsslein-Volhard, C. & Jürgens, G. *Wilhelm Roux Arch. dev. Biol.* **193,** 283-295 (1984).
5. Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193,** 296-307 (1984).
6. Lewis, E. B. *Nature* **276,** 565-570 (1978).
7. Kaufman, T. C., Lewis, R. & Wakimoto, B. *Genetics* **94,** 115-133 (1980).
8. Wakimoto, B. T. & Kaufman, T. C. *Devl Biol.* **81,** (1981).
9. Ouweneel, W. H. *Adv. Genet.* **18,** 179-205 (1976).
10. Wakimoto, B. T., Turner, F. R. & Kaufman, T. C. *Devl Biol.* **102,** 147-172 (1985).
11. Hafen, E., Levine, M., Garber, R. L. & Gehring, W. J. *EMBO J.* **2,** 617-623 (1983).
12. Akam, M. E. *EMBO J.* **2,** 2075-2084 (1983).
13. Hafen, E., Kuroiwa, A. & Gehring, W. J. *Cell* **37,** 833-841 (1984).
14. Martinez-Arias, A. & Lawrence, P. A. *Nature* **313,** 639-642 (1985).
15. Poole, S., Kauvar, L. M., Drees, B. & Kornberg, T. *Cell* **40,** 37-43 (1985).
16. Fjose, A., McGinnis, W. & Gehring, W. J. *Nature* **313,** 284-289 (1985).
17. Knipple, D. C., Seifert, E., Rosenberg, U. B., Preiss, A. & Jäckle, H. *Nature* **317,** 40-44 (1985).
18. Kilchher, F., Baumgartner, S., Bopp, D., Frei, E. & Noll, M. *Nature* **321,** 493-499 (1986).
19. Struhl, G. & White, R. H. *Cell* **43,** 507-519 (1985).
20. White, R. A. H. & Wilcox, M. *Nature* **318,** 563-567 (1985).
21. Harding, K., Wedeen, C., McGinnis, W. & Levine, M. *Science* **229,** 1236-1242 (1985).
22. Wedeen, C., Harding, K. & Levine, M. *Cell* **44,** 739-748 (1986).
23. Ingham, P. W., Howard, K. R. & Ish-Horowicz, D. *Nature* **318,** 439-445 (1985).
24. Howard, K. & Ingham, P. *Cell* **44,** 949-957 (1986).
25. Carroll, S. B. & Scott, M. P. *Cell* **45,** 113-126 (1986).
26. Ingham, P. W., Ish-Horowicz, D. & Howard, K. R. *EMBO J.* **5,** 1659-1665 (1986).
27. Akam, M. & Martinez-Arias, A. *EMBO J.* **4,** 1989-1700 (1985).
28. Meinhardt, H. *J. Cell Sci.* **4,** 357-381 (1986).
29. Preiss *et al. Nature* **317,** 27-32 (1985).
30. Rosenberg, U. B. *et al. Nature* **319,** 336-339 (1986).
31. Lehmann, R. & Nüsslein-Volhard, C. *Devl. Biol.* (in the press).
32. Wieschaus, E. *et al. Devl. Biol.* **104,** 172-186 (1984).
33. Kemphues, K. J. *et al., Genetics* **105,** 345-356 (1983).

Location of a δ-subunit region determining ion transport through the acetylcholine receptor channel

Keiji Imoto*‡, Christoph Methfessel*, Bert Sakmann*, Masayoshi Mishina†, Yasuo Mori†, Takashi Konno†, Kazuhiko Fukuda†, Masaaki Kurasaki†, Hideaki Bujo†, Yoshihiko Fujita† & Shosaku Numa†

* Max-Planck-Institut für Biophysikalische Chemie,
D-3400 Göttingen, FRG
† Departments of Medical Chemistry and Molecular Genetics,
Kyoto University Faculty of Medicine, Kyoto 606, Japan

The combination of complementary DNA expression and single-channel current analysis provides a powerful tool for studying the structure–function relationship of the nicotinic acetylcholine receptor (AChR) (refs 1–5). We have previously shown that AChR channels consisting of subunits from different species, expressed in the surface membrane of *Xenopus* oocytes, can be used to relate functional properties to individual subunits[4]. Here we report that, in extracellular solution of low divalent cation concentration, the bovine AChR channel has a smaller conductance than the *Torpedo* AChR channel. Replacement of the δ-subunit of the *Torpedo* AChR by the bovine δ-subunit makes the channel conductance similar to that of the bovine AChR channel. To locate the region in the δ-subunit responsible for this difference, we have constructed chimaeric δ-subunit cDNAs with different combinations of the *Torpedo* and bovine counterparts. The conductances of AChR channels containing chimaeric δ-subunits suggest that a region comprising the putative transmembrane segment M2 and the adjacent bend portion between segments M2 and M3 is involved in determining the rate of ion transport through the open channel.

The *Torpedo* and bovine AChRs were expressed in *Xenopus laevis* oocytes by microinjection of the respective α-, β-, γ- and δ-subunit-specific messenger RNAs synthesized by transcription *in vitro* of the cloned cDNAs (refs 2, 4, 6). Figure 1*a* compares the sizes of acetylcholine(ACh)-activated elementary inward currents mediated by the *Torpedo* (round symbols) and bovine

---

‡ Permanent address: Departments of Medical Chemistry and Molecular Genetics, Kyoto University Faculty of Medicine, Kyoto 606, Japan.

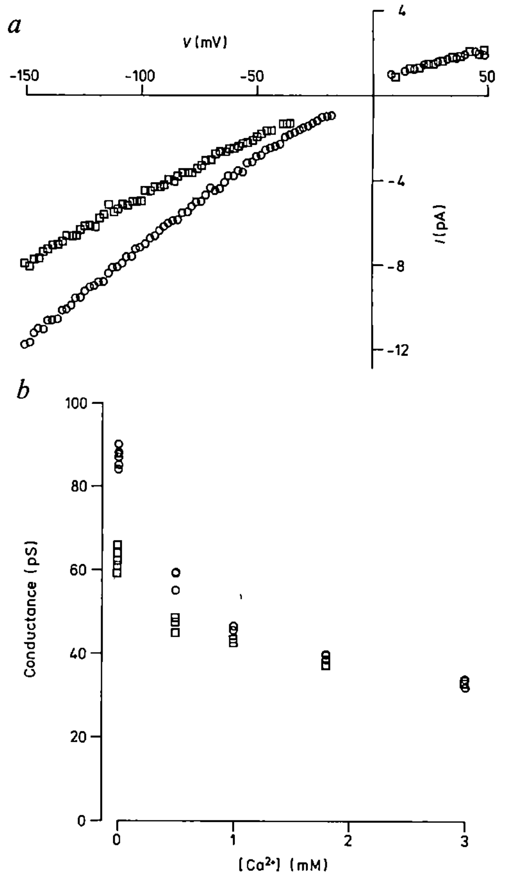


Fig. 1 Conductance properties of the *Torpedo* and bovine AChR channels produced in *Xenopus* oocytes *a*, Single-channel *I*-*V* relations of (○) *Torpedo* and (□) bovine AChR channels produced in *Xenopus* oocytes injected with the respective α-, β-, γ- and δ-subunit-specific mRNAs All single-channel *I*-*V* relations were measured at 12 ± 1 °C in cell-attached patches The pipette was filled with 30 mM KCl, 70 mM KF, 1 mM $MgCl_2$, 10 mM EGTA and 10 mM HEPES (*p*H 7.2), the concentration of free Mg^{2+} is calculated to be less than 1 μM, using a stability constant for MgEGTA of 1.6×10^5 M^{-1} This solution is referred to as low divalent cation extracellular solution (LDCS) Each point represents the mean amplitude of at least four elementary currents measured during either rampwise or stepwise changes of the voltage applied to the patch pipette In experiments with LDCS in the pipette, *I*-*V* relations were sometimes slightly bent in the inward direction at potentials more negative than −150 mV Conductances were measured as slope conductances between −40 mV and −150 mV membrane potential by linear regression The conductances of the *Torpedo* and bovine AChR channels illustrated are 88 pS ($r^2 = 0.997$) and 61 pS ($r^2 = 0.995$), respectively In experiments with low K^+ pipette solution (10 mM KCl, 5 mM EGTA and 10 mM HEPES (*p*H 7.2)), the slope conductances were measured at membrane potentials between −90 mV and −200 mV *b*, Ca^{2+} concentration dependence of single-channel K^+ conductances Slope conductances measured between −40 mV and −150 mV membrane potential are shown as a function of Ca^{2+} concentration in the pipette LDCS as in *a* In experiments with different Ca^{2+} concentrations, $CaCl_2$ was added to 100 mM KCl and 10 mM HEPES (*p*H 7.2) to give the desired Ca^{2+} concentration Each point represents measurement of the slope conductance in one patch ○, *Torpedo* AChR channels, □, bovine AChR channels

Methods The *Torpedo* and bovine AChR subunit-specific mRNAs were synthesized *in vitro* as described previously[2,4,6], except that the transcription was primed[19] with the cap dinucleotide G(5′)ppp(5′)G (0.5 mM) and that the bovine δ-subunit-specific mRNA was synthesized using *Eco*RI-cleaved pSPcδ2 as template (for construction, see Fig 3 legend) *Xenopus laevis* oocytes were injected with specified combinations of the *Torpedo* and bovine AChR subunit-specific mRNAs and incubated for 3-5 days[4] Single-channel current measurements were made from cell-attached membrane patches as described[20] The oocyte membrane potential was measured by an intracellular microelectrode and added to the pipette potential to give the patch transmembrane potential The potential across the patch was voltage clamped via the patch pipette to the desired command potential In each experiment, the current step sizes were measured during rampwise or stepwise changes of the patch membrane potential between −175 mV and 75 mV Each *I*-*V* relation is based on 300-1,400 amplitude determinations in the voltage range between −150 mV and 50 mV Current records were sampled at 20 kHz after low-pass filtering at 2 or 3 kHz (−3 dB) and analysed by a semiautomatic procedure The patch membrane potential was sampled at 0.6 kHz Single-channel current amplitudes within 2 mV intervals were pooled

(square symbols) AChR channels as measured in cell-attached patches with a potassium-rich pipette solution at low (<1 μM) divalent cation concentration The single-channel current-voltage (*I*-*V*) relations show that the K^+ conductance of the *Torpedo* AChR channel (87 ± 2 pS, mean ± s.d.), measured as slope conductance between −40 mV and −150 mV membrane potential, is larger than that of the bovine AChR channel (62 ± 2 pS) (Fig 3*b*) When the Ca^{2+} concentration in the pipette is increased, the conductances of both the *Torpedo* and bovine AChR channels decrease and the two tend to equalize (Fig 1*b*) At Ca^{2+} concentrations ≥ 1.8 mM, both channels are equally conductive A decrease in the conductances of the *Torpedo* and bovine AChR channels was observed in the presence of added extracellular Ca^{2+} with Na^+, Cs^+ or Li^+ as the permeant cation The reduction in single-channel conductance was also observed when the extracellular divalent cation was Mg^{2+}, Sr^{2+} or Ba^{2+} For both channels the effect of Mg^{2+} was more pronounced than that of the other divalent cations The K^+ conductances at low extracellular divalent cation concentration represent high concentration limits of these conductances since varying the K^+ concentration in the pipette between 10 mM and 100 mM altered the channel conductance of both the *Torpedo* and the bovine AChR by <10%

One explanation for the dependence of the AChR channel conductance on extracellular divalent cations is that it reflects changes in the effective negative surface charges on the receptor[7,8] At low divalent cation concentration, permeant cations would be concentrated at the mouth of the channel as compared with the bulk solution This concentration effect could cause ion saturation of the channel even when the cation concentration in the bulk solution is low (10 mM) Other mechanisms, such as blocking of monovalent cation flow by divalent cations, probably play a role at higher divalent cation concentrations[9] We have used the difference in conductance observed at low divalent cation concentration in an attempt to identify those parts of the AChR molecule important for forming the channel

To determine whether a particular subunit is responsible for this difference in conductance between the *Torpedo* and bovine AChR channels, we constructed hybrid AChRs by injecting *Xenopus* oocytes with the bovine α-subunit-specific mRNA and

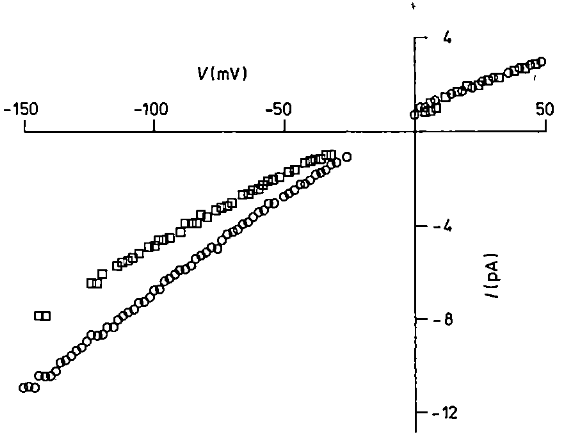


Fig 2 Single-channel *I*-*V* relations of α-hybrid (○) and δ-hybrid (□) AChR channels The hybrid AChR channels were produced in *Xenopus* oocytes by injection of the specified combinations of the *Torpedo* and bovine subunit-specific mRNAs Experimental details as in Fig 1*a*, pipettes filled with LDCS The slope conductances of the α-hybrid and δ-hybrid AChR channels shown are 84 pS ($r^2 = 0.998$) and 64 pS ($r^2 = 0.996$), respectively

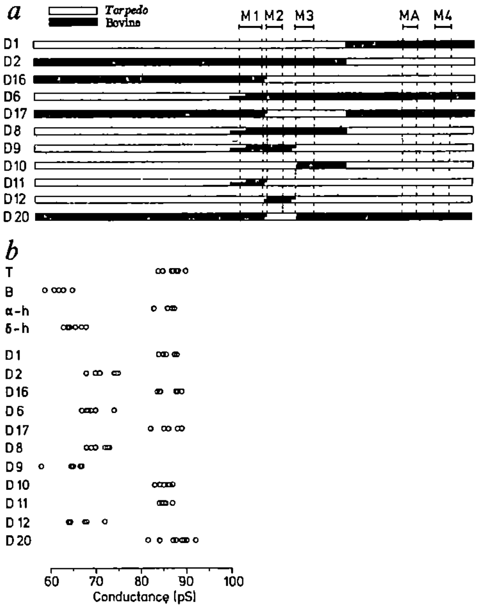


Fig 3 Location of a δ-subunit region determining the rate of ion transport through the AChR channel *a*, Structures of chimaeric δ-subunits composed of *Torpedo*[21] and bovine[15] sequences The structures of chimaeric δ-subunits, including the amino-terminal signal sequence, are diagrammatically shown such that the hydrophobic segments M1–M4 and the amphipathic segment MA (refs 2, 10–14) are aligned[15], the fact that the bovine δ-subunit is smaller by six amino acids than the *Torpedo* δ-subunit is ignored Open boxes, amino-acid sequences derived from the *Torpedo* δ-subunit, filled boxes, those from the bovine δ-subunit Junctional sequences common to the two δ-subunits are indicated by both boxes The compositions of the individual chimaeric δ-subunits are as follows (T and B denote the *Torpedo* and the bovine δ-subunit, respectively and the numbers in parentheses indicate amino-acid numbers[15,21], the junctional sequences common to the two δ-subunits are represented by amino acid numbers of the *Torpedo* δ-subunit) D1, T(−21–350) and B(350–495), D2, B(−21–350) and T(352–501), D16, B(−21–252) and T(254–501), D6, T(−21–231) and B(231–495), D17, B(−21–252), T(254–350) and B(350–495), D8, T(−21–231), B(231–350) and T(352–501), D9, T(−21–231), B(231–284) and T(286–501), D10, T(−21–292), B(292–350) and T(352–501), D11, T(−21–231), B(231–252) and T(254–501), D12, T(−21–256), B(256–284) and T(286–501), D20, B(−21–252), T(254–292) and B(292–495) *b*, Single-channel conductances of the *Torpedo* AChR (T), bovine AChR (B), α-hybrid AChR (α-h), δ-hybrid AChR (δ-h) and AChRs with chimaeric δ-subunits The AChRs with chimaeric δ-subunits were produced in *Xenopus* oocytes by injection of the respective chimaeric δ-subunit-specific mRNAs together with the *Torpedo* α-, β- and γ-subunit-specific mRNAs Each point represents measurement of slope conductance in one patch Experimental details as in Fig 1*a*, pipettes filled with LDCS The mean conductances of the individual AChR channels are as follows *Torpedo* AChR, 87±2 pS (mean±s d , $n=7$), bovine AChR, 62±2 pS ($n=5$), α-hybrid AChR, 86±2 pS ($n=4$), δ-hybrid AChR, 65±2 pS ($n=6$), AChR-D1, 86±2 pS ($n=5$), AChR-D2, 72±3 pS ($n=6$), AChR-D16, 86±2 pS ($n=6$), AChR-D6, 69±3 pS ($n=6$), AChR-D17, 86±3 pS ($n=5$), AChR-D8, 71±2 pS ($n=6$), AChR-D9, 64±4 pS ($n=5$), AChR-D10, 85±1 pS ($n=6$), AChR-D11, 85±1 pS ($n=5$), AChR-D12, 67±3 pS ($n=5$), AChR-D20, 88±3 pS ($n=8$)

Methods The 1 113-base-pair (bp) *Sma*I fragment from pSPcδ (ref 4) was ligated with the synthetic *Hin*dIII linker 5′-ACAAGCTTGT-3′, cleaved with *Hin*dIII and *Acc*I, and the resulting 423-bp fragment was isolated The 5 0-kilobase (kb) *Hin*dIII fragment from pSPcδ was treated with T4 DNA polymerase in the presence of the four deoxyribonucleoside triphosphates, ligated with the synthetic *Eco*RI linker 5′-GGAATTCC-3′ and cleaved with *Eco*RI and *Acc*I The resulting 1 5-kb fragment was ligated with the 423-bp fragment described above and the 3 0-kb *Hin*dIII–*Eco*RI fragment from pSP64 (ref 22) to yield pSPcδ2 The ~830-bp *Sma*I–*Eco*RI fragment from pSPcδ2, the 4 2-kb *Bam*HI–*Eco*RI fragment from pSPδ (ref 2) and the synthetic oligodeoxyribonucleotide

```
5 -GATCCTCCATATGTCGCGAGCTGATGAGAGCGAGCAGCCGGAT-3
    3 -GAGGTATACAGCGCTCGACTACTCTCGCTCGTCGGCCTA-5
```

(prepared with an automatic DNA synthesizer, Applied Biosystems) were ligated to yield pSPD1 The 4 1-kb *Xma*I–*Eco*RI fragment from pSPcδ2, the ~710-bp *Eco*RI–*Alu*I fragment isolated from the ~770-bp *Eco*RI–*Bam*HI fragment from pSPδ and

```
5 -CCGGGCAGAACGATTTGAAG-3
    3 -CGTCTTGCTAAACTTC-5
```

were ligated to yield pSPD2 The 1 1-kb *Hha*I–*Eco*RI fragment from pSPδ, the 3 7-kb *Eco*RI–*Aat*II fragment from pSPcδ2 and

```
      5 -CACCTTCTACCTCATTATCCGGCG-3
3 -TGCAGTGGAAGATGGAGTAATAGGCC-5
```

were ligated to yield pSPD3 The 4 1-kb *Eco*RI–*Sca*I fragment from pSPδ, the ~980-bp *Nco*I–*Eco*RI fragment from pSPcδ2 and

```
5 -ACCTGCTCTTTGGCATGGTGCTGGTGAC-3
3 -TGGACGAGAAACCGTACCACGACCACTGGTAC-5
```

were ligated to yield pSPD4 The 3 4-kb *Acc*I–*Eco*RI and 492-bp *Acc*I–*Nco*I fragments from pSPcδ2, the ~900-bp *Sca*I–*Eco*RI fragment from pSPδ and

```
5 -CATGGCCATACCTCTGATTGGTAAGT-3
    3 -CGGTATGGAGACTAACCATTCA-5
```

were ligated to yield pSPD5 The 678-bp *Sac*I–*Hpa*II fragment from pSPδ and

```
5 -CGGATAAATTCCCTAACGGCACCAATTACCAGGACGT-3
  3 -CTATTTAAGGGATTGCCGTGGTTAATGGTCC-5
```

were ligated and cleaved with *Sac*I and *Aat*II The resulting 715-bp fragment was ligated with the 3 1-kb *Sac*I–*Eco*RI fragment from pSPδ and the 1 2-kb *Aat*II–*Eco*RI fragment from pSPcδ2 to yield pSPD6 The ~870-bp *Bam*HI–*Eco*RI fragment from pSPD1 and the 4 1-kb *Eco*RI–*Bam*HI fragment from pSPD3 were ligated to yield pSPD7 The ~730-bp *Xma*I–*Eco*RI fragment from pSPD2 and the 4 2-kb *Xma*I–*Eco*RI fragment from pSPD6 were ligated to yield pSPD8 The 3 8-kb *Eco*RI–*Aat*II fragment from pSPD6 and the 1 1-kb *Aat*II–*Eco*RI fragment from pSPD5 were ligated to yield pSPD9 The ~730-bp *Xma*I–*Eco*RI fragment from pSPD2 and the 4 2-kb *Eco*RI–XmaI fragment from pSPD4 were ligated to yield pSPD10 The 253-bp *Dde*I–*Bgl*I fragment isolated from the 850-bp *Dde*I fragment from pSPδ was ligated with the 2 1-kb *Bgl*I fragment from pSPδ The resulting 2 3-kb fragment was ligated with the 2 4-kb *Bgl*I–*Hin*dIII fragment from pSPδ, the 180-bp *Hin*dIII–*Sau*3AI fragment isolated from the 228-bp *Hin*dIII–*Nco*I fragment from pSPD6 and

```
5 -GATCAACCTGGTCTTCTACTTGCCGGCCGACTGTGGTGAGAAAAT-
    3 -TTGGACCAGAAGATGAACGGCCGGCTGACACCACTCTTTTA-

GAGCACGGCTATTTCTGTCTTGCTTGC-3
CTCGTGCCGATAAAGACAGAACGAACGAGT-5
```

to yield pSOD[1] The 4 7-kb *Pst*I–*Bgl*II fragment from pSPδ, the 210-bp *Bgl*II–*Nco*I fragment from pSPD9 and

```
    5 -GAGAGTGGTGAGAAGACATC-3
3 -ACGTCTCTCACCACTCTTCTGTAGGTAC-5
```

were ligated to yield pSPD12 The 326-bp *Acc*I–*Fok*I fragment from pSPD5 and the 328-bp *Fok*I–*Bam*HI fragment from pSPD11 were ligated and cleaved with *Acc*I and *Bam*HI The resulting 654-bp fragment was ligated with the 4 0-kb *Bam*HI–*Pst*I and 150-bp *Pst*I–*Acc*I fragments from pSPD5 to yield pSPD16 The 326-bp *Acc*I–*Fok*I fragment from pSPD5 and the 328-bp *Fok*I–*Bam*HI fragment from pSPD11 were ligated and cleaved with *Acc*I and *Bam*HI The resulting 654-bp *Acc*I–*Bam*HI fragment was ligated with the 3 9-kb *Bam*HI–*Hin*dIII and 423-bp *Hin*dIII–*Acc*I fragments from pSPD7 to yield pSPD17 The 4 7-kb *Nco*I–*Aat*II fragment from pSPcδ2, the 153-bp *Aat*II–*Dde*I fragment isolated from the 363-bp *Aat*II–*Bam*HI fragment from pSPD11 and the 102-bp *Dde*I–*Nco*I fragment isolated from the 146-bp *Pst*I–*Nco*I fragment from pSPD4 were ligated to yield pSPD20 Those regions of the constructed plasmids derived from the synthetic oligodeoxyribonucleotides were sequenced[23] to ensure the constructions The chimaeric δ-subunit-specific mRNAs were synthesized as described in the legend to Fig 1, using *Eco*RI-cleaved plasmids carrying the respective cDNAs as templates

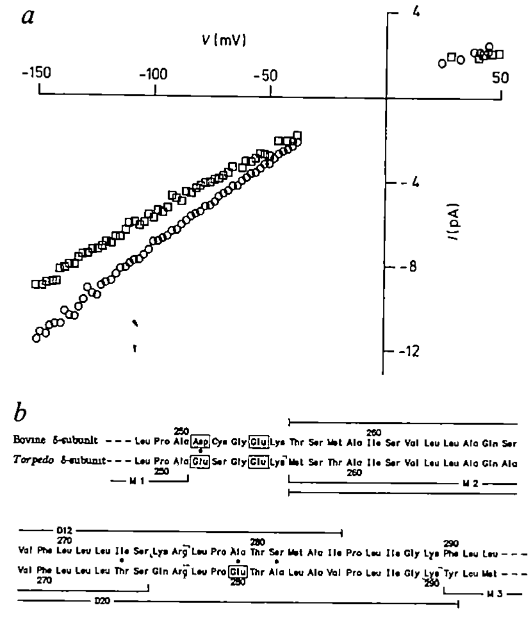


Fig 4 *a*, Single-channel *I*–*V* relations of the (□) AChR-D12 and (○) AChR-D20 channels Experimental details as in Fig 1*a*, pipettes filled with LDCS The slope conductances of the AChR-D12 and AChR-D20 channels shown are 65 pS ($r^2 = 0\ 994$) and 85 pS ($r^2 = 0\ 997$), respectively *b*, Comparison of the amino-acid sequences of the bovine (top) and *Torpedo* (bottom) AChR δ-subunits in the M2 and adjacent regions The bovine and *Torpedo* sequences are aligned[15] and the bovine sequence in the D12 subunit and the *Torpedo* sequence in the D20 subunit are indicated Asterisks, amino-acid differences between the two sequences, negatively charged and positively charged residues are enclosed with solid and dashed lines, respectively Amino-acid numbers[15 21] are given and the positions of segments M1, M2 and M3 are indicated, note that the amino-acid numbers of the two δ-subunits do not coincide with the alignment because of a gap present upstream

the *Torpedo* β-, γ- and δ-subunit-specific mRNAs (α-hybrid AChR, ref 4) or with the bovine δ-subunit-specific mRNA and the *Torpedo* α-, β- and γ-subunit-specific mRNAs (δ-hybrid AChR, ref 4) Figure 2 shows the single-channel *I*–*V* relations of the α-hybrid (round symbols) and δ-hybrid (square symbols) AChR channels in low divalent cation extracellular solution The K$^+$ conductance of the α-hybrid AChR channel is 86±2 pS (Fig 3*b*) This is similar to the conductance of the parental *Torpedo* AChR channel, suggesting that a difference in the α-subunit is not responsible for the larger conductance of the *Torpedo* AChR channel In contrast, the K$^+$ conductance of the δ-hybrid AChR channel is 65±2 pS (Fig 3*b*), being similar to that of the bovine AChR channel Thus, the difference in K$^+$ conductance between the *Torpedo* and bovine AChR channels can be ascribed, at least partly, to a difference in their δ-subunits The contributions of the β- and γ-subunits were not assessed because of the very low channel activities of the β-hybrid and γ-hybrid AChRs expressed in *Xenopus* oocytes[4]

To localize that part of the δ-subunit molecule responsible for the difference in the *Torpedo* and δ-hybrid AChR channel conductances, we prepared cDNA constructs encoding chimaeric δ-subunits in which portions of the *Torpedo* δ-subunit were systematically replaced by homologous regions of the bovine δ-subunit (Fig 3*a*) The mRNA derived from each of these cDNA templates, together with the *Torpedo* α-, β- and γ-subunit-specific mRNAs, was injected into *Xenopus* oocytes to produce AChRs with the different chimaeric δ-subunits Figure 3*b* shows graphically the K$^+$ conductances, measured in low divalent cation solution, of these AChR channels with the different chimaeric δ-subunits The conductances observed fall into one of two broad classes of lower, bovine-like (64–72 pS) or higher, *Torpedo*-like (85–88 pS) mean conductance

The conductance of the AChR-D1 channel (86±2 pS) and that of the AChR-D2 channel (72±3 pS) suggest that the portion of the δ-subunit comprising the amino-terminal extracellular region and the putative transmembrane segments M1, M2 and M3 (refs 10–14) is responsible for the different K$^+$ conductances of the *Torpedo* and δ-hybrid AChR channels Comparison of the conductances of the AChR-D16 channel (86±2 pS) and the AChR-D6 channel (69±3 pS) further suggests that the critical region lies in the portion containing segments M2 and M3 To refine this mapping, we tested AChRs with δ-subunit chimaeras in which only small segments had been exchanged

The conductances of the AChR-D17 channel (86±3 pS) and the AChR-D8 channel (71±2 pS) are in agreement with the above localization Furthermore, the conductances of the AChR-D9, AChR-D10 and AChR-D11 channels (64±4 pS, 85±1 pS and 85±1 pS, respectively) exclude a major involvement of segment M3 as well as of segment M1 Finally, the conductance of the AChR-D12 channel (67±3 pS), in which only the segment composed of amino-acid residues 256–284 of the bovine δ-subunit[15] has been substituted for the *Torpedo* sequence, is similar to that of the δ-hybrid AChR channel (see Fig 4*a*, square symbols and Fig 4*b*) In accord with this result is the conductance of the AChR-D20 channel (88±3 pS), which is similar to that of the *Torpedo* AChR channel (see Fig 4*a*, round symbols) The D20 subunit is nearly a 'mirror image' of the D12 subunit in that only the segment composed of amino-acid residues 256–290 of the bovine δ-subunit has been replaced by the *Torpedo* sequence (Fig 4*b*) The AChR-D12 and the AChR-D20 channel showed a Ca^{2+} dependence of their conductance similar to that of the bovine and the *Torpedo* AChR channel, respectively

From these results, we conclude that the region of the δ-subunit comprising segment M2 and the adjacent bend portion between segments M2 and M3 is principally responsible for the difference in K$^+$ conductance between the *Torpedo* and bovine AChR channels In line with this conclusion is the finding that replacement of the *Torpedo* α-subunit by the bovine α-subunit, both of which contain an identical amino-acid sequence in this region, does not alter the channel conductance Thus, our results suggest that the M2 and bend region represents an important determinant of the rate of ion transport through the AChR channel Recently it has been reported that segment M2 contributes to the binding site of noncompetitive antagonists[16–18]

In the bend portion between the putative transmembrane segments M2 and M3, the *Torpedo* δ-subunit has one negative and one positive charge, whereas the bovine δ-subunit has two positive charges, the lysine residue immediately preceding segment M3 is not taken into account because the five amino-acid residues preceding segment M3 are common to the two δ-subunits (Fig 4*b*) All subunits of both the *Torpedo* and bovine AChRs contain negatively charged residues in this bend portion[10,15], except the bovine δ-subunit A hypothesis consistent with our results is that the negatively charged residues in the bend portion of the AChR subunits are located near the mouth of the channel, attracting permeant cations toward the channel One effect of substituting the D12 chimaeric subunit for the *Torpedo* δ-subunit would be to reduce the number of negative charges in the mouth region Alternatively, the M2 and bend region may constitute part of the channel wall and its amino-acid sequence may partly determine the energy profile experienced by permeating cations in the channel

We thank Drs S DeRiemer, F Conti and E Neher for reading the manuscript This investigation was supported in part by the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases

Received 13 October, accepted 14 November 1986

1 Mishina, M *et al* *Nature* **307**, 604–608 (1984)
2 Mishina, M *et al* *Nature* **313**, 364–369 (1985)

3 Sakmann, B & Neher, E *A Rev Physiol* **46**, 455-472 (1984)
4 Sakmann, B *et al Nature* **318**, 538-543 (1985)
5 Mishina, M *et al Nature* **321**, 406-411 (1986)
6 Takai, T *et al Nature* **315**, 761-764 (1985)
7 McLaughlin, S' G 'A, Szabo, G & Eisenman, G *J gen Physiol* **58**, 667-687 (1971)
8 Hille, B, Woodhull, A M & Shapiro, B I *Phil Trans R Soc Lond B* **270**, 301-318 (1975)
9 Marty, A *J Physiol, Paris* **76**, 523-527 (1980)
10 Noda, M *et al Nature* **302**, 528-532 (1983)
11 Claudio, T, Ballivet, M, Patrick, J & Heinemann, S *Proc natn Acad Sci USA* **80**, 1111-1115 (1983)
12 Devillers Thiery, A, Giraudat, J, Bentaboulet, M & Changeux, J P *Proc natn Acad Sci USA* **80**, 2067-2071 (1983)
13 Guy, H. R. *Biophys J* **45**, 249-261 (1984)
14 Finer Moore, J & Stroud, R M *Proc natn, Acad Sci USA* **81**, 155-159 (1984)
15 Kubo, T *et al Eur J Biochem* **149**, 5-13 (1985)
16 Giraudat, J, Dennis, M, Heidmann, T, Chang, J-Y & Changeux, J-P *Proc natn Acad Sci USA* **83**, 2719-2723 (1986)
17 Oberthur, W, Muhn, P, Baumann, H, Lottspeich, F, Wittmann Liebold, B & Hucho, F *EMBO J* **5**, 1815-1819 (1986)
18 Hucho, F, Oberthür, W & Lottspeich, F *FEBS Lett* **205**, 137-142 (1986)
19 Konarska, M M, Padgett, R A & Sharp, P A *Cell* **38**, 731-736 (1984)
20 Methfessel, C *et al Pflügers Arch ges Physiol* (in the press)
21 Noda, M *et al Nature* **301**, 251-255 (1983)
22 Melton, D A *et al Nucleic Acids Res* **12**, 7035-7056 (1984)
23 Sanger, F, Nicklen, S & Coulson, A R *Proc natn Acad Sci USA* **74**, 5463-5467 (1977)

Abnormal tyrosine phosphorylation on T-cell receptor in lymphoproliferative disorders

Lawrence E. Samelson, Wendy F. Davidson*, Herbert C. Morse III† & Richard D. Klausner

Cell Biology and Metabolism Branch, National Institute of Child Health, and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA
* Laboratory of Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA
† Laboratory of Immunopathology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

The study of human autoimmune diseases has benefited greatly from analysis of animal models. Mice that are homozygous for either the *lpr* (lymphoproliferation)[1] or *gld* (generalized lymphoproliferative disease)[2] mutant genes develop a disease characterized by massive lymphadenopathy and autoantibody formation. With age, the lymphoid organs in these mice are replaced with a greatly expanded population of abnormal lymphocytes[3–5]. Recent work has shown that these cells are likely to be in the T-cell lineage. They rearrange and transcribe the genes for the α and β subunits of the T-cell receptor (TCR) and a third, T-cell receptor-like gene, T_γ[6,7] As determined by immunofluorescence with anti-receptor antibodies the cells also express TCR on the cell surface. The murine T-cell receptor consists of the α and β chains, derived from the rearranged α and β genes, in non-covalent association with seven other chains; the δ chain, of relative molecular mass (M_r) 26,000 (26K), the ε chain (25K), a glycosylated 21K chain (gp21) which is probably the homologue of the γ chain of T3 (CD3), a 16K homodimer (ζ) and a 21K dimer (p21)[8–10]. This multichain complex is thought to be the murine analogue of the human T3 complex. After activation of normal T cells by antigen or lectin, p21 is phosphorylated on tyrosine residues and gp21 is phosphorylated on serine residues. In contrast, in the *gld* and *lpr* cells, p21 is phosphorylated even in the absence of antigen or lectin, whereas gp21 is not phosphorylated.

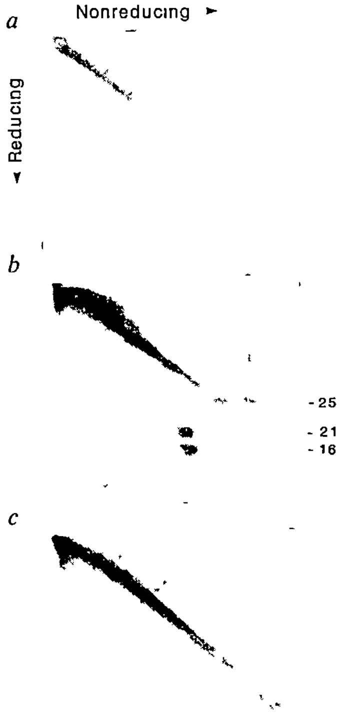


Fig 1 Two-dimensional diagonal SDS-polyacrylamide gel electrophoresis (PAGE) of phosphorylated normal and *lpr* T cell antigen receptors Purified Ly2⁻, L3T4⁻, lymph node cells lacking surface immunoglobulin from C3H-*lpr* animals (*a*, *b*) and purified T cells from normal C3H spleens and lymph nodes (*c*) were ^{32}P-labelled and solubilized in 0 5% Triton X-100 Immunoprecipitations were performed with preimmune rabbit serum (*a*) or with rabbit anti-δ chain serum[9] (*b*, *c*)

Methods The abnormal C3H-*lpr* and *gld* T cells (L3T4$^-$, Ly2$^-$) were separated from normal L3T4$^+$ and Ly-2$^+$ T cells by panning as described[7] C3H +/+ T cells were isolated from a mixture of spleen and lymph node cells by incubating 5×10^7 cells on anti-immunoglobulin coated plates for 45 min at 20 °C as described previously[7] Surface-immunoglobulin-negative cells were recovered by gently washing the plates five times After concentration by centrifugation the cells were applied to a fresh set of anti-Ig coated plates for a second cycle of purification The cells were washed in phosphate-free RPMI1640, resuspended in this medium with 10% dialysed fetal calf serum at $2 4\times10^7$ cells per ml and incubated with 0 8 mCi ml^{-1} [^{32}P]orthophosphate for 1 h at 37 °C The cells were then washed and solubilized as previously described[10] Immunoprecipitations were performed with 20 μl of preimmune serum or anti-δ serum adsorbed to protein A-Sepharose The immunoprecipitates were washed and eluted as described[10] The eluents were electrophoresed in the first dimension (left to right) without reduction in a 13% SDS-PAGE tube gel The tubes were equilibrated in 0 5% dithiothreitol, 0 1% SDS, 0 125 M Tris *p*H 6 8 for 2 h at room temperature before electrophoresis in a 13% SDS-PAGE slab gel in the second dimension Gels were treated and autoradiographed as described[10] Each panel represents immunoprecipitation of lysates from 6×10^7 cells Molecular weight markers are shown The gel was exposed to film for 48 h at −70 °C

The autosomal recessive mutations *lpr* and *gld* are non-allelic genes, yet the abnormal cells that populate the lymph nodes of mice homozygous for either mutation are indistinguishable by a wide variety of criteria We have suggested that the mutations may affect different components of a common metabolic pathway of major importance to T-cell differentiation and function[4] The abnormal cells of *lpr* and *gld* mice appear to express a normal complement of clonally determined α-β heterodimers at the cell surface but fail to respond to simulation with alloantigens[7], raising the possibility that the T-cell receptor is involved in the mutation Therefore, the antigen receptors on the abnormal T cells of *lpr* and *gld* mice were isolated and compared to those of normal peripheral T cells The cells were purified by plate separation techniques, surface radio-iodinated and solubilized in non-ionic detergents The antigen receptors were isolated with an antibody that binds the murine antigen receptor δ chain and is capable of immunoprecipitating the entire complex[9] All the components of the receptor are specifically co-immunoprecipitated with the anti-δ chain reagent from both

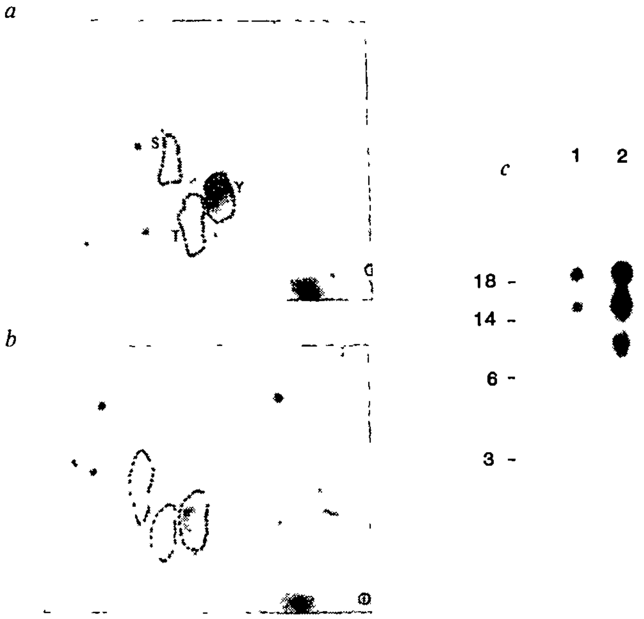


Fig 2 *a*, *b*, Phosphoamino acid analysis of phosphorylated antigen receptor subunits from abnormal *gld* cells Phosphorylated 21K chains (*a*) or 16K chains (*b*) from C3H-*gld* cells were analysed (*c*) Phosphopeptide mapping of antigen receptor subunits by limited proteolysis Phosphorylated p21 subunits from antigen activated 2B4 cells (lane 1) and MRL/*lpr* cells (lane2) were subjected to limited proteolysis with staphylcoccal V8 protease

Methods Phosphoamino acid content of proteins was determined using a modification of the method of Cooper *et al*[17] The abnormal cells from C3H-*gld* and C3H-*lpr* lymph nodes were isolated, labeled with ^{32}P and solubilized as described in Fig 1 Labelled proteins were separated on two dimensional diagonal gels and detected by autoradiography as described in Fig 1 Gel regions containing individual proteins were excised and hydrolysed with 2 ml of 5 7 M HCl at 110 °C for 1 h After addition of 4 ml of water, the hydrolysates were centrifuged at 500*g* for 10 min, and the supernatants were lyophilized The residues were dissolved in 10 ml of water, to which was added Dowex AG1-X8 (400 μl of 50% v/v), and phosphoamino acid standards (1 μl of 1 mg ml^{-1} stock solutions) The *p*H of the solutions was adjusted to 7 5 to 8 5 with 10 mM NH$_4$OH Solutions were tumbled with the ion exchange resin at room temperature for 4 h and centrifuged The supernatant was reapplied to a second batch of resin and tumbled overnight The two sets of beads were washed with water and eluted with four sequential 0 5 ml applications of 0 1 M HCl Eluents were dried under vacuum, pooled in 1 ml of water, redried under vacuum, and stored at 4 °C Samples were mixed with phosphoamino acid standards and analysed by two dimensional high voltage electrophoresis on thin layer plates as described[16] Electrophoresis was performed at *p*H 1 9, at 900 V for 1 5 h in the first (vertical) dimension and at *p*H 3 5, at 900 V for 40 min in the second (horizontal) dimension Gels were exposed to film for 12 days The pigeon cytochrome *c* hybridoma 2B4 ($2\ 5\times10^8$) was ^{32}P labelled and activated as described elsewhere[11] The antigen receptor was isolated with the clonotypic monoclonal, A2B4-2[18] MRL/*lpr* cells (10^9) were phosphorylated and the antigen receptor was isolated as described in Fig 1 Phosphorylated components were isolated on two-dimensional diagonal gels, spots containing the p21 chains were cut out, and the polypeptides were subjected to proteolysis with V8 protease in the gel as described[19] A 15% acrylamide gel was used to resolve the phosphopeptides *a*, 48 h exposure, *b*, 16 h exposure

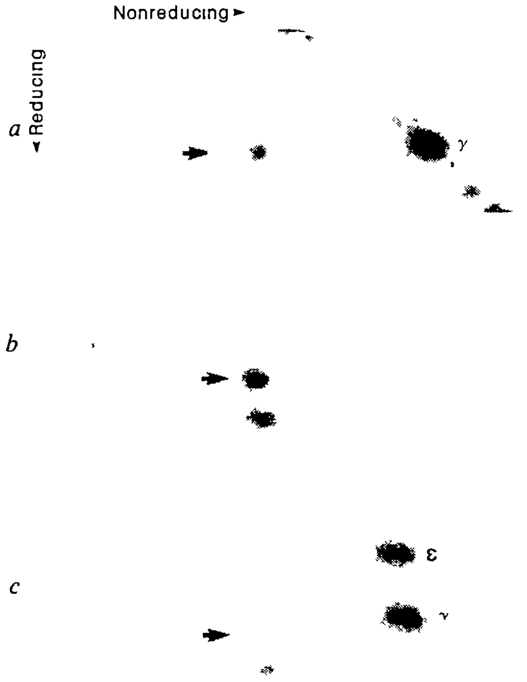


Fig 3 Two dimensional SDS-PAGE analysis of phosphorylated T cells Phosphorylated antigen receptors from *a*, antigen-activated 2B4 cells, *b*, C3H-*gld* cells, *c*, PMA treated C3H-*gld* cells were electrophoresed on 2D diagonal gels Arrow, phosphorylated p21 subunit

Methods The pigeon cytochrome *c* specific hybridoma 2B4 was ^{32}P-labelled and activated by antigen and the antigen receptor was isolated with the clonotypic monoclonal, A2B4-2 as described in Fig 2*a* C3H-*gld* cells were isolated as described in Fig 1*b*, *c* After phosphate labelling half of the *gld* cells were treated for 30 min with TPA at 100 ng ml^{-1} (*c*) Receptors from *gld* cells were isolated with the anti-δ chain antibody Electrophoresis and autoradiography were as described (Fig 1)

normal and abnormal T cell populations (data not shown) Several subunits of the T-cell antigen receptor are phosphorylated during T-cell activation[8,10-12] The 21K γ glycoprotein and, to a lesser degree, the 25K ε chain are phosphorylated on serine residues, and the p21 structure is phosphorylated on tyrosine residues within minutes of either antigen or mitogen presentation to T cell hybridomas Several other receptors are also phosphorylated on tyrosine upon activation, tyrosine phosphorylation is otherwise very rare[13] The serine phosphorylation is presumably due to the activation of protein kinase C and can be mimicked by the addition of TPA (tetradecanoyl phorbol 13-acetate) The kinase responsible for tyrosine phosphorylation has not been identified, but p21 phosphorylation is associated only with presentation of specific antigen or mitogenic lectin There is little, if any, phosphorylation of receptor chains without activation in normal cells Surprisingly, in contrast to normal T cells, the *lpr* (Fig 1) and *gld* T cell (Fig 3*b*) receptors were heavily phosphorylated in the absence of any added stimulus Labelled phosphates are added in the course of a one hour incubation This rapid turnover implies constitutive activation of both a kinase and a phosphatase Receptors isolated from T cells of animals with either mutation are constitutively phosphorylated on both 21 and 16K chains The 16K chain may be a proteolytic cleavage of the 21K chain but their exact relationship is still unclear As these bands are vertically aligned below the diagonal on this two dimensional gel, it is likely that they are linked under non-reducing conditions, they probably form a dimer in the receptor We have shown that specific antigen activation of T cell hybridomas results in phosphorylation of the p21 component of the antigen receptor on tyrosine residues[8,12], so we isolated the 21K and 16K proteins from the *gld* and *lpr* cells to determine which aminoacid was phosphorylated Both chains from the two mutant strains were phosphorylated on tyrosine (Fig 2*a*, *b*) Partial phosphopeptide mappings of p21 from *lpr* cells and an antigen-stimulated T-cell hybridoma with V8 protease (Fig 2*c*) and cyanogen bromide (not shown) are consistent with an identical site of phosphorylation The differences between the tyrosine phosphorylations on abnormal and normal T cells cannot be explained by enhanced phosphate exchange The turnover rate of the p21 tyrosine phosphate in *gld* cells is the same (~20 min) as that seen in activated normal T cells (data not shown) In normal T cells there is no tyrosine phosphorylation on p21 in the absence of stimulation The constitutive phosphorylation in the *gld* and *lpr* cells must be due to abnormalities in the tyrosine kinase pathway

In our previous studies, we found that whenever antigen or mitogen activates T cells, two parallel phosphorylation pathways are activated resulting in the serine phosphorylation of gp21 and the tyrosine phosphorylation of p21 A relatively fixed relationship between the intensities of these two phosphorylations is found over a wide range of antigen doses An example (Fig 3*a*) shows that the intensity of the serine phosphorylation is greater than that of tyrosine phosphorylated p21 in a T cell hybridoma Normal splenic T cells stimulated for 1 hour or 40 hours with the mitogen concanavalin A also show both serine and tyrosine phosphorylations (data not shown) The relative absence of the 16K phosphoprotein in the 2B4 cells appears to be peculiar to this hybridoma as mitogen-stimulated splenic T cells also show the 21K and 16K tyrosine-phosphorylated doublet In contrast, the *gld* (Fig 3*b*) and *lpr* (Fig 1*b*) cells demonstrate no detectable gp21 serine phosphorylation despite high levels of p21 tyrosine phosphorylation The addition of mitogens to *lpr* and *gld* cells failed to enhance the tyrosine phosphorylation and did not result in gp21 serine phosphorylation, again in contrast to normal T cells This is consistent with the failure to proliferate in response to stimulation with concanavalin A We examined whether the absence of gp21 serine phosphorylation was due to lack of activatable protein kinase C Addition of TPA to normal T cells or T-cell hybridomas bypasses the receptor and directly activates this kinase resulting in serine, but not tyrosine phosphorylation of antigen receptor subunits We therefore added TPA to ^{32}P-labelled *gld* cells (Fig 3*c*) This results in the expected gp21 phosphorylation and enhanced ε phosphorylation indicating that protein kinase C is both present and capable of phosphorylating the *gld* gp21 subunit Interestingly, TPA treatment seems to inhibit phosphorylation of p21

The constitutive tyrosine phosphorylation in the absence of serine phosphorylation distinguishes the abnormal cells of *gld* and *lpr* from bulk populations of resting splenic T cells, mitogen activated T-cell blasts or antigen-activated T-cell hybridomas, but the fact that the abnormal T cells in these disorders are phenotypically different from the T cells to which they were compared makes the interpretation of these differences difficult We do not know whether the T cells in these animals are aberrant lineages unique to the disease, or an abnormal expansion of a rare T cells subset present in normal animals If these cells do exist in normal mice, they might show the same phosphorylation phenomena reported here The recent report of 'double negative' T cells in the periphery of normal individuals may be due to just such populations[13,14] The fact that the T cell antigen receptor plays such a crucial role in the growth and differentiation of T cells makes understanding its function in abnormal T cells essential The observation that a biochemical pathway identified with receptor activation is constitutively activated in these abnormal T cells may offer clues to the pathogenesis of the disease processes A defect in the coupling mechanism between the receptor and kinase, or abnormalities in intracellular kinase regulation may be the manifestation of the mutant genes The clear role of tyrosine kinases in both normal and abnormal growth and differentiation suggests parallels to this system Current views about the *erb*B oncogene, for example, postulate that its transforming activity is due to the functioning of a constitutively active tyrosine kinase no longer under the regulation of a growth factor hormone[15,16] Two observations reported here suggest a similar dissociation from normal ligand regulation (1) the constitutive tyrosine phosphorylation in the absence of any apparent ligand, and (2) the lack of receptor serine phosphorylation that characterizes normal ligand induced receptor-activation Pursuing the implications and consequences of these unusual phosphorylations may elucidate mechanisms involved in both normal and abnormal T cell function

We thank H Luong for technical assistance and K Perry for preparation of the manuscript

Received 7 August, accepted 24 October 1986

1 Murphy, E D & Roths, J B in *Genetic Control of Autoimmune Disease* (eds Rose, N R, Bigazzi, P E & Warner, N L) 207-220 (Elsevier, New York, 1978)
2 Roths, J B, Murphy, E D & Eicher, E M *J exp Med* **159**, 1-20 (1984)
3 Morse, H C III *et al. J Immun* **129**, 2612-2615 (1982)
4 Davidson, W F, Holmes, R L, Roths, J B & Morse, H C III *Proc. natn. Acad. Sci. U S A.* **82**, 1219-1223 (1985)
5 Davignon, J L *et al. J Immun* **135**, 2423-2428 (1985)
6 Nemazee D A, Studer, S, Steinmetz, M, Dembic, Z & Kiefer, M *Eur J Immun* **15**, 760-764 (1985)
7 Davidson, W F, Dumont, F J, Bedigian, H G, Fowlkes, B J & Morse, H C III *J Immun* **136**, 4075-4084 (1986)
8 Samelson L E, Patel, M D, Weissman, A M, Harford, J B & Klausner, R D *Cell* **46**, 1083-1090 (1986)
9 Samelson L E, Weissman, A, Robey, F A, Berkower, I & Klausner, R D *J Immun.* **137**, 3254-3258
10 Samelson L E, Harford, J B & Klausner, R D *Cell* **43**, 223-231 (1985)
11 Samelson, L E, Harford, J B, Schwartz, R H & Klausner, R D *Proc. natn Acad Sci. U S A* **82**, 1969-1973 (1985)
12 Patel M Samelson L E & Klausner R. D (in preparation)
13 Brenner, M B *et al. Nature* **322**, 145-149 (1986)
14 Lanier, L L, Ruitenberg, J J & Phillips, J H *J exp Med* **164**, 339-344 (1986)
15 Hunter, T & Cooper, J A *A Rev Biochem.* **54**, 897-930 (1985)
16 Downward, J *et al. Nature* **307**, 521-527 (1984)
17 Cooper, J A, Sefton, B M & Hunter, T *Meth Enzym* **99**, 387-402 (1983)
18 Samelson, L E, Germain, R N & Schwartz, R H *Proc. natn. Acad. Sci U S A.* **80**, 6972-6976 (1983)
19 Cleveland, D W Fischer, S G, Kirchner, M W & Laemmli, U K. *J biol Chem.* **252**, 1102-1106 (1977)

Polymorphism of human Ia antigens generated by reciprocal intergenic exchange between two DR β loci

Shaokee Wu, Thomas L. Saunders & Fritz H. Bach

Immunobiology Research Center, Box 724 Mayo Memorial Building,
University of Minnesota, Minneapolis, Minnesota 55455, USA

Class II molecules encoded by the human major histocompatibility complex (MHC) are involved in regulating T-cell response to antigens[1–3]. The mechanisms for generating polymorphism in products of the MHC have been studied extensively for both the murine H–2 and the human HLA complex. Such studies indicate that point mutations plus selection have a major role in the generation of polymorphisms of class I and class II MHC genes[4–8]. However, a non-reciprocal gene conversion mechanism has been proposed to explain several examples of clustered sequence variation in MHC genes[9–14]. In all these examples, the proposed gene conversion event is unidirectional; that is, one of the two interacting genes acts as sequence donor and the other as sequence recipient. No examples of potential reciprocal genetic exchange (as occurs in the fungal system[15,16]), in which the two interacting genes act as both donor and recipient of gene fragments, have been found in the MHC system or in other multigene families of higher organisms. We sequenced two different HLA-DR β complementary DNAs from each of two different cells all expressing the same serologically defined determinant (DR2) but different T-cell-recognized (Dw) specificities (Dw12 and MN2). Sequence comparisons of these four cDNA clones (and two DR β amino-acid sequences from the DR2–Dw2 subtype) suggest that new coding sequences for DR β molecules in the DR2 haplotypes are potentially generated by reciprocal intergenic exchange

Table 1 lists six amino-acid sequences for DR β genes from three different DR2 subtypes (DR2-Dw2, DR2-Dw12 and DR2-MN2)[17] (For simplicity only amino acids that differ between any two sequences are given, see ref 18 for a complete listing

Table 1 Grouping and comparison of amino-acid sequences of six DR β-chains from HLA-DR2 individuals

			Amino-acid position																													
			First domain														Second domain											CP	TM		CY
Group	β-chain	Dw	6	9	11	13	28	30	31	37	38	47	67	70	71	85	86	96	104	105	108	110	120	133	135	142	150	157	191	203	222	231
I	Dw2-A	Dw2	R	Q	D	Y	H	D	I	D	L	Y	F	D	R	V	G	E	A	R	T	E	—	R	S	V	—	—	—	—	—	—
	DHO-7	Dw12	—	—	—	—	—	—	—	—	—	Y	F	D	R	V	G	E	A	R	T	Q	N	R	S	V	N	T	Q	V	K	H
	MN2-2	MN2	C	Q	D	Y	H	G	I	N	V	Y	I	Q	A	A	V	E	A	R	T	Q	N	R	G	V	N	I	Q	I	K	H
II	Dw2-B	Dw2	R	W	P	R	D	Y	F	S	V	F	I	Q	A	V	V	Q	S	K	P	Q	S	L	G	M	D	T	R	V	R	Q
	DHO-8	Dw12	R	W	P	R	D	Y	F	S	V	F	I	Q	A	V	G	Q	S	K	P	Q	S	L	G	M	N	T	R	V	R	Q
	MN2-61	MN2	R	W	P	R	D	Y	F	S	V	Y	F	D	R	V	G	Q	S	K	P	Q	S	L	G	M	N	T	R	V	R	Q

The table lists differences in the amino-acid sequences of six DR β-chains from three different HLA-DR2-positive individuals Positions not shown in the table are identical in all six DR β-chains (see ref 18 for a complete listing of shared amino-acid residues) The Dw2-A and Dw2-B sequences are from protein sequencing of two DR β-chains of a DR2-Dw2 homozygous cell line[18] The other four sequences are deduced from cDNA clones obtained from cDNA libraries constructed in our laboratory These four cDNAs encode DR β molecules as revealed by restriction mapping (especially characteristic internal *Pst*I sites) and comparison with available DR β sequences[6 7] DHO (DR2-Dw12) is a homozygous typing cell (HTC) line[17] The MN2-2 and MN2-61 sequences are from an individual typed as DR2-MN2/DR4-Dw4 MN2-2 and MN2-61 are considered to be DR β sequences from the MN2 haplotype because (1) their sequences are different from the two published DR β genes of the DR4-Dw4 haplotype[35 41], (2) they have greater homology to other DR2-derived DR β sequences than to DR4 DR β sequences, (3) we have sequenced another DR β_1 cDNA from the same cell line and found it to be identical to other DR4-DW4 DR β_1 sequences[35 41] (S W *et al*, unpublished observation) cDNA was synthesized using oligo(dT) to prime the first strand and nascent messenger RNA was nicked with RNase H to prime the second strand synthesis[22] cDNA was 'blunt-ended' with DNA polymerase and tailed with dCTP using terminal deoxynucleotidyl transferase cDNA was annealed with *Pst*I-cut, d*G*-tailed single-stranded plasmid vector pUC119 (ref 42) and introduced into *E coli* strain MV1193 Each library, which contained about 50,000 colonies, was screened with a 790-bp *Sct*I/*Hin*dIII fragment of a DR β cDNA clone[43] labelled with ^{32}P by nick translation This probe contains almost the entire coding region from this clone and under the hybridization conditions used, reacts most strongly with DR β cDNA but also reacts weakly with DQ β and DP β cDNAs The inserts were sequenced directly on pUC119 or truncated subclones were generated by the method of Dale *et al*[44] The second orientation was sequenced from inserts subcloned into plasmid vector pUC118 or M13 mp18 Sequencing was done in both orientations using the dideoxy chain termination method[45] with the M13 universal sequencing primer[46] or oligonucleotide primers corresponding to conserved DR β sequences The DHO-7 clone contains a truncated insert The available sequence of DHO-7 starts at amino-acid[45] Amino-acid residues of interest are boxed Residues not available for comparison are marked with a dash CP refers to the connecting peptide, TM to the transmembrane region and CY to the cytoplasmic tail of the DR β molecule The total length of the DR β-chain, not including the signal peptide, comprises 237 amino acids

Table 2 χ-Like sequences in the immunoglobulin–MHC supergene family

Species	Gene	Altered nucleotide/consecutive nucleotides*	χ-Like sequence	Refs
Murine	MHC-class II A_β^{bm12}	3/14	GCTGGGG	11, 12
Murine	MHC-class I K^{bm1}	7/13	GCTGGTG	9, 10
Human	MHC-class II DR3-DRβI	7/34	GCTGGGG	14
Human	MHC-class II DR2-DR β	7/18	GCTGGGG	This paper
Human	MHC-class I B27k	4/14	GCTGGTG	13
Murine	IgD V gene	15/164	GCCTGGGG(CTGA)GCTGGTG(AA)GCCTGGGG	32

The χ sequence is GCTGGTGG or its complement or its duplex[26] The χ-like sequence for the immunoglobulin switch region is GCTGGTG(G) (ref 28), for Ti-plasmid GCTGGTGG and GCTGTTGG (ref 29), for a core sequence in human 'minisatellite' DNAs GGAGGTGGGCAGGA$^{A}_{G}$G (ref 30), and for murine class II E_β gene $(GGCA)_{25-37}$ (L Hood, personal communication)

* The number of altered nucleotides/the number of consecutive nucleotides from the first to last alteration

of shared amino-acid residues) The six sequences can be classified unambiguously into two groups by sequence homology, they presumably represent the two expressed DR β molecules in the DR2 haplotype[19,20] The group I DR β molecules are apparently more polymorphic than the group II DR β molecules Of greatest importance are amino acids 67–71 (Table 1) These five residues are shared by MN2-2 in group I and two sequences in group II (Dw2-B and DHO-8), also, MN2-61 in group II shares these five residues with two sequences in group I (Dw2-A and DHO-7) Corresponding cDNA nucleotide sequences of these DR β molecules (Fig 1) show that DHO-7 (group I) and MN2-61 (group II) share all 18 nucleotides coding for residues 67–72, as do DHO-8 (group II) and MN2-2 (group I) These two 18-nucleotide sequences differ at 7 positions, resulting in 3 amino-acid substitutions out of 6 It is extremely unlikely that seven nucleotide differences are due simply to the accumulation of sequential point mutations[21], especially as the third base of each codon is the same for each amino acid The simplest explanation of this finding is that there has been reciprocal genetic exchange of this segment between the two non-allelic expressed DR β genes

The non-reciprocal gene conversion events implicated in generating the A_β^{bm12} allele (murine)[11,12], DR3 haplotype (human)[14] and various DR4-Dw subtypes[14,22,23] involve the same third hypervariable region in the first domain of the class II gene This hypervariable region apparently specifies an epitope for alloreactivity[11,22–24] and restricted recognition[25] The finding that this functionally important segment of the class II gene is frequently involved in gene conversion-like events prompted us to look for a unique sequence which may promote genetic recombination[15,16 26,27] These include the χ sequence in λ phage and *Escherichia coli* (5'-GCTGGTGG-3' or its complement or its duplex)[26] and the repeated T G and C G dinucleotides implicated in a gene conversion event in the human globin genes[27]

χ-Like sequences have been recognized in DNA which undergo rearrangement or recombination (refs 28–30 and L Hood, personal communication) (Table 2) A χ-like sequence (5'-GCTGGGG-3') is found 36 base pairs (bp) upstream from the 18-bp 'interconverted' hypervariable segment within a region of complete sequence homology from amino-acid residue 48 to 66 among all four DR β cDNAs that we have sequenced (from the second exon) (Fig 1) The amino-acid sequences for Dw2-A and Dw2-B (ref 18) (Table 1) are consistent with such a nucleotide sequence

There are two other χ-like sequences in the DR2 β cDNAs, one near amino acid 1 (from the first exon) (Fig 1) and another near residue 140 (5'-GCTGGGG-3' for both MN2-2 and DHO-7, and 5'-GCTGGGA-3' for both MN2-61 and DHO-8, these

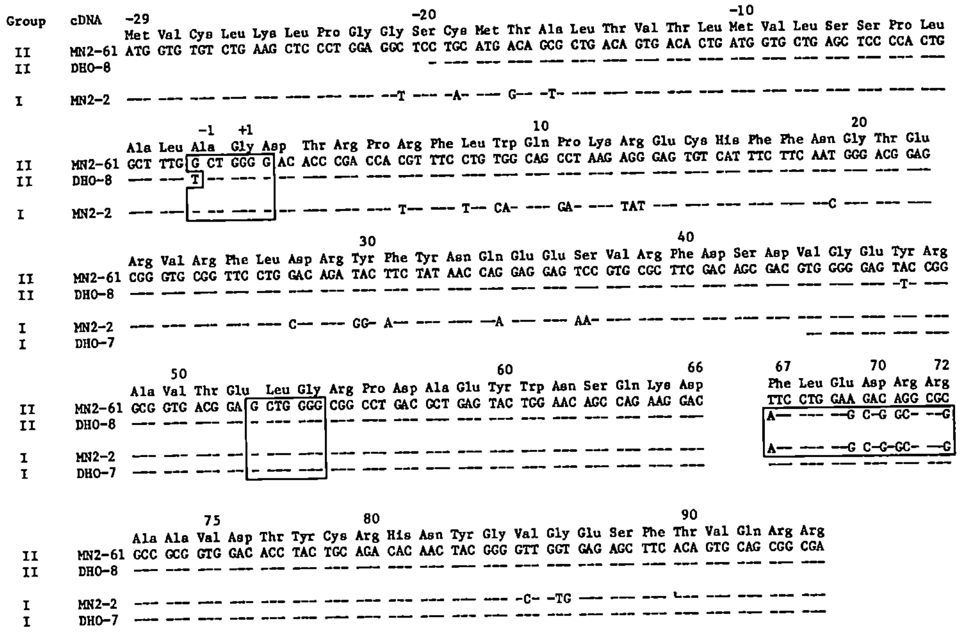

Group cDNA -29 -20 -10
Met Val Cys Leu Lys Leu Pro Gly Gly Ser Cys Met Thr Ala Leu Thr Val Thr Leu Met Val Leu Ser Ser Pro Leu
II MN2-61 ATG GTG TGT CTG AAG CTC CCT GGA GGC TCC TGC ATG ACA GCG CTG ACA GTG ACA CTG ATG GTG CTG AGC TCC CCA CTG
II DHO-8 - --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
I MN2-2 --- --- --- --- --- --- --- --- --T --- -A- --- G-- -T- --- --- --- --- --- --- --- --- --- --- --- ---
-1 +1 10 20
Ala Leu Ala Gly Asp Thr Arg Pro Arg Phe Leu Trp Gln Pro Lys Arg Glu Cys His Phe Phe Asn Gly Thr Glu
II MN2-61 GCT TTG G CT GGG G AC ACC CGA CCA CGT TTC CTG TGG CAG CCT AAG AGG GAG TGT CAT TTC TTC AAT GGG ACG GAG
II DHO-8 --- --- T -- --- - -- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
I MN2-2 --- --- - -- --- - -- --- --- --- T-- --- T-- CA- --- GA- --- TAT --- --- --- --- --- --C --- --- ---
30 40
Arg Val Arg Phe Leu Asp Arg Tyr Phe Tyr Asn Gln Glu Glu Ser Val Arg Phe Asp Ser Asp Val Gly Glu Tyr Arg
II MN2-61 CGG GTG CGG TTC CTG GAC AGA TAC TTC TAT AAC CAG GAG GAG TCC GTG CGC TTC GAC AGC GAC GTG GGG GAG TAC CGG
II DHO-8 --- -T- ---
I MN2-2 --- --- --- --- --- C-- --- GG- A-- --- --- --A --- --- AA- --- --- --- --- --- --- --- --- --- --- ---
I DHO-7 --- --- --- --- ---
50 60 66 67 70 72
Ala Val Thr Glu Leu Gly Arg Pro Asp Ala Glu Tyr Trp Asn Ser Gln Lys Asp Phe Leu Glu Asp Arg Arg
II MN2-61 GCG GTG ACG GA G CTG GGG CGG CCT GAC GCT GAG TAC TGG AAC AGC CAG AAG GAC TTC CTG GAA GAC AGG CGC
II DHO-8 --- --- --- -- - --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --- --G C-G GC- --G
I MN2-2 --- --- --- -- - --- --- --- --- --- --- --- --- --- --- --- --- --- --- A-- --- --G C-G-GC- --G
I DHO-7 --- --- --- -- - ---
75 80 90
Ala Ala Val Asp Thr Tyr Cys Arg His Asn Tyr Gly Val Gly Glu Ser Phe Thr Val Gln Arg Arg
II MN2-61 GCC GCG GTG GAC ACC TAC TGC AGA CAC AAC TAC GGG GTT GGT GAG AGC TTC ACA GTG CAG CGG CGA
II DHO-8 ---
I MN2-2 --- --- --- --- --- --- --- --- --- --- --- --- -C- -TG --- --- --- L-- --- --- --- ---
I DHO-7 ---

Fig 1 Nucleotide sequences of four cDNA clones encoding HLA-DR β molecules from two different HLA-DR2 subtypes The sequences shown include the leader sequence (from residue −29 to −1) and the first domain (residues 1-94) DHO-7 and DHO-8 cDNAs are from DR2-Dw12, and MN2-2 and MN2-61 from DR2-MN2 The deduced amino-acid sequence for the DR β cDNA clone MN2-61 is given above the MN2-61 nucleotide sequence Dashes in the sequences indicate identity with the MN2-61 sequence Numbers correspond to the amino-acid sequence of the DR β molecule with a leader sequence 29 amino acids long The 5'-GCTGGGG-3' sequence is boxed due to its similarity to the χ sequence (5'-GCTGGTGG-3') of λ phage, a signal for generalized recombination in *E coli*[26]

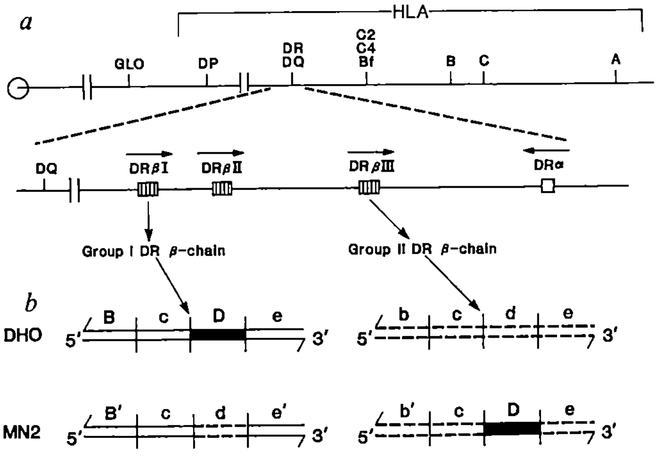


Fig 2 Genetic map of the human MHC[34,35,47] and schematic representation of the reciprocal genetic exchange in the second exon (first domain) of the non-allelic, expressed DR β genes *a*, The human MHC is located on the short arm of chromosome 6 The relative position of DP genes, DQ genes, the DR α gene and DR β genes on the map was determined by pulse-field electrophoresis mapping[47] and by genomic cloning and sequencing[34,35] The DRβII gene in both DR3 and DR4 haplotypes is a pseudogene[34,38] It remains to be proven whether the HLA-DR2 haplotype also has a DR β pseudogene In this figure we suggest that the DRβI gene encodes the more polymorphic group I DR β molecules (Table 1) and that the DRβIII gene encodes the less polymorphic group II DR β molecules (Table 1), as in the DR3 and DR4 (refs 14, 35) haplotypes, but this remains to be conclusively demonstrated by genomic and/or amino-acid sequencing The circle represents the centromere, and arrows show the direction of transcription[34,35] *b*, For ease of illustration, the second exon (amino acids 5-94) of the DR β gene is divided into four segments B/b amino acids 5-47, c, 48-66, D/d, 67-72, and e, 73-94 B/B'/b/b', D/d, e/e' indicate sequence variation in that particular segment between alleles and/or loci Identical letters indicate sequence identity The χ-like sequence 5'-GCTGGGG-3' is found in the conserved c segment The D/d segment is the minimum 18-nucleotide sequence being potentially interconverted See Table 1 for amino-acid sequences and Fig 1 for nucleotide sequences

sequences are from the third exon) We also identified χ-like sequences in published sequences of other class I and class II MHC genes as well as immunoglobulin genes (which belong to the same supergene family as the MHC genes[31]) where gene conversion-like events have been proposed[9-14,32] (Table 2) The frequent observations of χ-like sequences upstream from the potentially 'converted' gene segments suggest that χ-like sequences may be associated with recombination hot spots involved in a gene conversion-like event in the immunoglobulin-MHC supergene family Promotion of recombination by χ-like sequences may be tested by using tetrad analysis of yeast recombination system[33]

The DR genomic organization of the DR2 haplotype is not known If we assume that it is similar to that of the DR3 (ref 34) and DR4 (ref 35) haplotypes (Fig 2*a*), the genetic events which lead to the formation of current DR β genes in different DR2 subtypes may be postulated as follows (1) gene duplication and deletion from a primordial gene of the immunoglobulin supergene family[7,31] to generate one DR α gene and three DR β genes (Fig 2*a*), (2) point mutations generating DR β gene polymorphism including the 18 bases encoding residues 67-72, (3) genetic interactions and recombination between two non-allelic DR β loci (probably the DRβI and DRβIII genes), leading to the exchange of residues 67-72 (Fig 2*b*) and (4) further point mutations Whether the third DR β gene (DRβII, which may or may not be a pseudogene for the DR2 haplotype) is involved in the recombination process is unknown

To explain the genetic interactions and recombination, we hypothesize that the intergenic exchange was initiated by the χ-like sequence and involved the two non-allelic DRβI and DRβIII genes on the same chromatid looping back on itself in a single genetic event with the reciprocal exchange of gene segments[36,37] Our model would be equally applicable if exchange occurred between sister chromatids or between chromosomes that have a high degree of sequence homology and carry the two 18-nucleotide sequences encoding residues 67-72 (Fig 2) However, we do not exclude the possibility that reciprocal intergenic exchange resulted from two temporally independent non-reciprocal gene conversion events, or that double crossing-over can account for our observation, given the fact that 24-74 bp of complete sequence homology are sufficient for genetic recombination (segments c and e in Figs 1 and 2)[38-40], although this is less likely given interference between closely linked genetic markers in crossing over

Gene conversion and reciprocal intergenic exchanges in MHC genes frequently involve hypervariable regions specifying functional epitopes and may allow the re-shuffling of functionally important epitopes to increase the combinatorial usage of a limited number of hypervariable regions and to create novel H-2 and HLA haplotypes

We thank M Madden, L Taaffe, D Lu, G Gordon and C P Liu for technical help, members of the Immunobiology Research Center for discussion, N Andresen and C Greenberg for manuscript preparation, B K Birshtein, L E Hood, P Powers, C M Radding and O Smithies for stimulating discussions and J Vierra and B Wold for technical advice This is paper 438 from the Immunobiology Research Center This work was supported in part by NIH grants AI 17687, AI 18326, AI 19007, AI 22682, clinical research grant 6-425 from the March of Dimes Birth Defects Foundation, and the Harry Kay Charitable Foundation

Received 18 May, accepted 9 September 1986

1 Benecerraf, B *Science* **212**, 1229-1238 (1981)
2 Kaufman, J F, Auffray, C Korman, A J, Shackelford, D A & Strominger, J L *Cell* **36**, 1-13 (1984)
3 Bach, F H *Immun Today* **6**, 89-94 (1985)
4 Gustafsson K *et al EMBO J* **3**, 1655-1661 (1984)
5 Bell, J I, Denny, D W Jr & McDevitt H O *Immun Rev* **84**, 51-71 (1985)
6 Auffray, C & Strominger, J L *Adv hum Genet* **15**, 197-247 (1985)
7 Figueroa, F & Klein, J *Immun Today* **7**, 78-81 (1986)
8 Klein, J & Figueroa, F *Immun Today* **7**, 41-44 (1986)
9 Meller, A L Weiss, E H, Ramachandran, K & Flavell, R A *Nature* **306**, 792-795 (1983)
10 Nathenson, S G, Geliebter J, Pfaffenbach, G M & Zeff, R A *A Rev Immun* **4**, 471-502 (1986)
11 Mengle Gaw, L, Conner, S, McDevitt, H O & Fathman, C G *J exp Med* **160**, 1184-1194 (1984)
12 Denaro M Hammerling, U, Rask, L & Peterson, P A *EMBO J* **3**, 2029-2032 (1984)
13 Seemann, G H A, Rein, R S, Brown, C S & Ploegh H L *EMBO J* **5**, 547-552 (1986)
14 Gorski, J & Mach, B *Nature* **322**, 67-70 (1986)
15 Stahl F W *Genetic Recombination* (Freeman, San Francisco, 1979)
16 Whitehouse H L K *Genetic Recombination* (Wiley, New York, 1982)
17 Reinsmoen, N L, Layrisse, Z, Betuel, H, Freidel, A C & Bach, F H *Hum Immun* **11**, 105-115 (1984)
18 Kratzin, H *et al* in *HLA Class II Antigens* (eds Solheim B G, Moller, E & Ferrone, S) 49-70 (Springer, Berlin, 1986)
19 Sone, T *et al J Immun* **135**, 1288-1298 (1985)
20 Nepom G T *et al Hum Immun* **10**, 143-151 (1984)
21 Klein J *Transplantation* **38**, 327-329 (1984)
22 Cairns, J S *et al Nature* **317**, 166-168 (1985)
23 Gregerson, P K *et al Proc natn Acad Sci U S A* **83**, 2642-2646 (1986)
24 Beck B N Nelson P A & Fatham, C G *J exp Med* **157**, 1396-1404 (1983)
25 Hochman, P S & Huber, B T *J exp Med* **160**, 1925-1930 (1984)
26 Smith G R, Kunes S M, Schultz, D W, Taylor A & Triman, K L *Cell* **24**, 429-436 (1981)
27 Slightom, J L, Blechl, A E & Smithies, O *Cell* **21**, 627-638 (1980)
28 Kenter A L & Birshtein, B K *Nature* **293**, 402-404 (1981)
29 Yadav, N S, Vanderleyden, J, Bennett, D R, Barnes, W M & Chilton, M D *Proc natn Acad Sci U S A* **79**, 6322-6326 (1982)
30 Jeffreys, A J, Wilson, V & Thein, S L *Nature* **314**, 67-73 (1985)
31 Hood, L E, Kronenberg, M & Hunkapiller, T *Cell* **40**, 225-229 (1985)
32 Krawinkel, U, Zoebelein, G, Bruggemann, M, Radbruch A & Rajewsky, K *Proc natn Acad Sci U S A* **80**, 4997-5001 (1983)
33 Treco, D, Thomas, T & Arnheim, N *Molec cell Biol* **5**, 2029-2038 (1985)
34 Rollini, P Mach, B & Gorski, J *Proc natn Acad Sci U S A* **82**, 7197-7201 (1985)
35 Spies T Sorrentino, R, Boss, J M Okada, K & Strominger, J L *Proc natn Acad Sci U S A* **82**, 5165-5169 (1985)
36 Meselson, M S & Radding, C M *Proc natn Acad Sci U S A* **72**, 358-361 (1975)
37 Smithies O & Powers, P A *Phil Trans R Soc* **B312**, 291-302 (1986)
38 Singer B S Gold L, Gauss P & Doherty D H *Cell* **31**, 25-33 (1982)
39 Shaul, Y, Laub, O, Walker, M D & Rutter, W J *Proc natn Acad Sci U S A* **82**, 3781-3784 (1985)
40 Watt, V M, Ingles, C J, Urdea M S & Rutter, W J *Proc natn Acad Sci U S A* **82**, 4768-4772 (1985)
41 Curtsinger J M, Hilden, J M, Cairns, J S & Bach, F H *Proc natn Acad Sci U S A* (in the press)
42 Vierra, J & Messing, J *Meth Enzym* (in the press)
43 Long, E C, Wake C T Gorski, J & Mach, B *EMBO J* **2**, 389-394 (1983)
44 Dale, R, McClure, B & Houchins, J P *Plasmid* **13**, 13-40 (1985)
45 Sanger, F, Nicklen, S & Coulson A R *Proc natn Acad Sci U S A* **74**, 5463-5467 (1977)
46 Messing, J *Meth Enzym* **101**, 20-78 (1983)
47 Hardy, D, Bell, J I, Long, E Lindsten, T & McDevitt, H O *Abstr 6th int Congr Immun* 142 (1986)
48 Larhammar, D, Servenius, B, Rask, L & Peterson, P A *Proc natn Acad Sci U S A* **82**, 1475-1479 (1985)

Predominant use of a V_α gene segment in mouse T-cell receptors for cytochrome *c*

Astar Winoto*, James L. Urban, Nancy C. Lan, Joan Goverman, Leroy Hood & Daniel Hansburg†

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA
† Department of Pathology, Fox Chase Cancer Center,
7701 Burholme Avenue, Philadelphia, Pennsylvania 19111, USA

The T-cell receptor is a cell surface heterodimer consisting of an α and a β chain that binds foreign antigen in the context of a cell surface molecule encoded by the major histocompatibility complex (MHC), thus restricting the T-cell response to the surface of antigen presenting cells[1-4]. The variable (*V*) domain of the receptor binds antigen and MHC molecules and is composed of distinct regions encoded by separate gene elements—variable (V_α and V_β), diversity (D_β) and joining (J_α and J_β)—rearranged and joined during T-cell differentiation to generate contiguous V_α and V_β genes[5-11]. T-helper cells, which facilitate T and B cell responses, bind antigen in the context of a class II MHC molecule. The helper T-cell response to cytochrome *c* in mice is a well-defined model for studying the T-cell response to restricted antigen and MHC determinants. Only mice expressing certain class II molecules can respond to this antigen ($E^k_\alpha E^k_\beta$, $E^k_\alpha E^b_\beta$, $E^v_\alpha E^v_\beta$ and $E^k_\alpha E^s_\beta$)[12-14]. Most T cells appear to recognize the C-terminal peptide of cytochrome *c* (residues 81–104 in pigeon cytochrome *c*)[15-18]. We have raised helper T cells to pigeon cytochrome *c* or its C-terminal peptide analogues in four different MHC congenic strains of mice encoding each of the four responding class II molecules. We have isolated and sequenced seven V_α genes and six V_β genes and analysed seven additional helper T cells by Northern blot to compare the structure of the V_α and V_β gene segments with their antigen and MHC specificities. We have added five examples taken from the literature[19-22]. These data show that a single V_α gene segment is responsible for a large part of the response of mice to cytochrome *c* but there is no simple correlation of MHC restriction with gene segment use.

The junctional nucleotide sequences for the V_α and V_β genes are given in Figs 1 and 2, respectively, and data for all the T cells examined are summarized in Table 1 These data, with functional studies to be discussed subsequently, demonstrate several interesting points

The $V_\alpha 11$ 1 gene segment is the one most often used in the mouse helper T-cell response to cytochrome *c* (14 out of 19 T_H cells) The closely related $V_\alpha 11$ 2 gene segment (eight nucleotide substitutions out of 312) is used in the 15th T_H cell The other predominant gene segments are used far less frequently than $V_\alpha 11$ 1 ($J_\alpha 84$, 6 of 12, $V_\beta 3$, 7 of 19, $J_\beta 1$ 2, 7 of 12) Some of

---

* Present address Whitehead Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA

Table 1 Gene segment usage by T-cell receptors recognizing pigeon cytochrome *c*

		α chain		β chain	
T cells*	Haplotype	*V*†	*J*	*V*	*J*
AN6 2	$E^k_\alpha E^k_\beta$	11 1‡	84	3	1 2
V11 5	$E^k_\alpha E^b_\beta$	11 1	84	3	1 2
C F6	$E^k_\alpha E^k_\beta$	11 1	84	3	1 2
AN14 4	$E^k_\alpha E^k_\beta$	11 1	14 4	3	1 2
5C C7	$E^k_\alpha E^k_\beta$	11 1	C7	3	1 2
BC15 1	$E^k_\alpha E^k_\beta$	11 1	84	Y#	ND
BC37 5	$E^k_\alpha E^k_\beta$	11 1	37	14#	ND
B10	$E^k_\alpha E^k_\beta$	11 1	84	B10	2 1
4 C3	$E^k_\alpha E^k_\beta$	11 1	84	B10	2 1
V1 9 2	$E^k_\alpha E^b_\beta$	11 1	28	1	1 1
AR8 1	$E^v_\alpha E^v_\beta$	11 1§	ND	Y#	ND
V15 4	$E^k_\alpha E^b_\beta$	11§	ND	8 3	1 4
AP10 10	$E^k_\alpha E^s_\beta$	11§	ND	8#	ND
AR10 1	$E^v_\alpha E^v_\beta$	11§	ND	Y#	ND
2B4	$E^k_\alpha E^k_\beta$	11 2	2B4	3	2 5
AP11 2	$E^k_\alpha E^s_\beta$	4 3	11 2	3	1 2
AP15 2	$E^k_\alpha E^s_\beta$	X§	ND	1	1 2
AN4 4	$E^k_\alpha E^k_\beta$	X§	ND	8#	ND
AP7 10	$E^k_\alpha E^s_\beta$	X§	ND	8#	ND

*C F6, 5C C7, B10 and 4 C3 are cloned helper T-cell lines induced in B10 A mice immunized with pigeon cytochrome *c*[22] The rest of the cells are T-helper hybridomas generated as described[46] 2B4 has been reported previously[19 20] Hybridomas were made by fusing BW5147 lymphoma cells with cytochrome *c*-specific T-cell lines induced in four different congenic mouse strains B10 A ($E^k_\alpha E^k_\beta$, AN and BC hybridomas), B10 SM ($E^v_\alpha E^v_\beta$, AR hybridomas), B10 A (5R) ($E^k_\alpha E^b_\beta$; V hybridomas) and B10 S (9R) ($E^k_\alpha E^s_\beta$, AP hybridomas) The mice were immunized with either pigeon cytochrome *c* (2B4 and BC hybridomas) or synthetic C-terminal homologous peptides to cytochrome *c* Sp (AN hydridomas) or DASp (V, AP, AR hybridomas)[46] B10 A and B10A(5R) hybridomas respond to antigen presented on $E^k_\alpha E^k_\beta$ or $E^k_\alpha E^b_\beta$ antigen presenting cells (APC) and B10 S(9R) and B10 SM hybridomas respond to antigen on $E^k_\alpha E^s_\beta$ or $E^v_\alpha E^v_\beta$ APC[12]

† Gene segment use was determined by DNA sequence analysis (for methods see legend of Fig 1) unless otherwise indicated Nomenclature is essentially as described[21-23] with the exception of J_α14 4 and J_α11 2 which have not been reported previously (for sequence see legend of Fig 1) The two distinct members of the V_α11 subfamily are denoted V_α11 1 and V_α11 2, these two members have been previously referred to as V_α2B4 2 and V_α2B4 1, respectively[22] They differ by only eight nucleotides in the coding region and nine nucleotides in the 5′ flanking region (ref 22 and data not shown) The V_α gene of hybridoma AP11 2 is denoted V_α4 3 to distinguish it from the two published V_α sequences belonging to the V_α4 subfamily[23] V_β3 has also been referred to as V_β2B4[22]

‡ Boxes indicate predominant gene segment use

§ Determined by Northern blotting using gene segments from V1 9 2 and AP11 2 as probes Use of V_α11 1 by AR8 1 was determined by additional Southern genomic blotting using a V_α11 flanking region probe V_α other than V_α4 and V_α11 are denoted by 'X'

# Determined by Northern blotting using probes from the V_β2,3,4,5,8 and 14 subfamilies V_β other than those used as probes denoted by 'Y'

```
                 ┌────────V────────┐  ┌──N or D──┐  ┌────────J────────┐
B10     Vα11 1— TGTGCTGCTGAG                        GCAACTTCAAGTG— Jα84
AN6 2   Vα11 1—                     AG*                         — Jα84
VII 5   Vα11 1—                     AG*                         — Jα84
C F6    Vα11 1—                     C  *                        — Jα84
BC15 1  Vα11 1—                                                 — Jα84
4 C3    Vα11 1—                     T  *                        — Jα84
AN14 4  Vα11 1—                     G  +                 T  A  T— Jα14 4
5C C7   Vα11 1—                     G  +                 C  A  A— JαC7
BC37 5  Vα11 1—                     TT +            A  AATGCA   — Jα37
VI 9 2  Vα11 1—                    CAGG*                 TG  CA— Jα28
2B4     Vα11 2—                 CCTGAGG+     TC     GG  G  CA— Jα2B4
APII 2  Vα4 3 —                 CTGGAGA+                 TG  A  — Jα11 2
```

Fig 1 V_α junctional nucleotide sequences of T-cell receptors recognizing cytochrome *c* Left, names of T-helper hybridomas and clones analysed Sequences of the C F6, B10, 4 C3, 5C C7 T-cell clones[22] and of the 2B4 hybridoma[19 20] have been reported previously and are included for comparison Variable (*V*) and joining (*J*) gene segment sequences up- and downstream of those shown were found to be identical to previously published sequences[19-23] with the exception of V_α4 3, J_α14 4 and J_α11 2, which have not been reported previously The V_α4 3 gene segment belongs to the V_α4 subfamily (at least 75% homologous to two previously reported V_α4 sequences[23]) and differs from the V_α11 1 gene segment of hybridoma AN6 2 by 80% at the protein level and 61% at the DNA level The V_α4 3 sequence is

GGCTTCGTGGCTGTGATACTTCTCATACTTGGAAGGAC
CCACGGAGATTCCGTGACTCAAACAGAAGGCCCAGTGA
CCGTCTCAGAAAGCGAGTCCCTGATAATAAATTGCACGT
ATTCAGCCACAAGCATAGCTTACCCTAATCTTTTCTGG
TATGTTCGATATCCTGGAGAAGGTCTACAACTCCTCC
TGAAAGTCATTACGGCTGGCCAGAAGGGAAGCAGCA
GAGGGTTTGAAGCCACATACAATAAAGAAACCACCTC
CTTCCACTTGCAGAAAGCCTCAGTGCAAGAGTCAGA
CTCGGCTGTGTACTACTGTGCT

The J_α14 4 sequence is

CTTCTAATTACAACGTGCTTTACTTCGGATCTGGCACC
AAACTCACTGTAGAGCCAAGTAAGTAAGTTATTGTTG
TTACCCACCCATTATGCTAAAT

The J_α11.2 sequence is

CTGGAGGCTATAAAGTGGTCTTTGGAAGTGGGACTCGA
TTGCTGGTAAGCCCTG

V and *J* region boundaries were determined by comparison to the germline sequences of V_α11 1 (ref 22), J_α84 (ref 47) and J_α28 (A W, unpublished data) and to the TA37 cDNA sequence (J_α37, ref 23) Nucleotides which could not be assigned to either *V* or J gene segments are designated 'N or D' and represent either additional non-germline sequences (*) or could not be evaluated because germline sequences are not known (+) Dots indicate identity to the B10 sequence

Methods The V_α genes of T-cell clones C F6, B10, 4 C3 and 5C C7 (ref 22) and of hybridoma 2B4 (refs 19, 20) were isolated from complementary DNA libraries as described The V_α genes of hybridomas V1 9 2, V11 5 and AP11 2 were isolated from λgt10 cDNA libraries and subcloned into m13mp8 The rest of the hybridomas were isolated from limited genomic libraries constructed in λL47 1 or EMBL3 vectors and were subcloned into M13 mp18 and/or 19 Both strands of DNA were sequenced using the dideoxy chain termination method[48] with the M13 universal primer and a C_α-specific oligonucleotide (CA3) primer for the cDNA clones, and specific oligonucleotide primers for the genomic clones

this apparent degeneracy may reflect the diversity of antigens and/or MHC molecules used in immunization[22] (see Table 1), but the V_α11 1 protein sequence is clearly highly favoured in the construction of a specific antigen binding site for cytochrome *c*

Some nucleotides in the V_α junctional region are not encoded by either germline V_α or J_α gene segments (Fig 1) Comparison of the known germline J_α84 and J_α28 sequences with those given in Fig 1 allows us to determine where these J_α sequences end in the 5′ direction Although the germline sequence of V_α11 1 is not known, clearly four of the five different sequences denoted by asterisks in Fig 1 must represent extra nucleotides (the fifth could be the germline V_α11 1 sequence) Two explanations are possible First, the extra nucleotides could represent contributions from D_α gene segments This explanation seems unlikely because the number of added nucleotides is small (1-4) and no other evidence for D_α gene segments has yet been reported[23] The second possibility is that N-region diversification, the addition of random nucleotides in the junctional region to either of the joined gene segments[24,25], may occur It has been proposed that N-region diversification is caused by the enzyme terminal deoxynucleotidyltransferase (TdT)[26], which is present in T cells

		V	D	J	
			Dβ1 1 GGGACAGGGGGC		
AP11 2	Vβ3	—TGTGCCAGCAGTCTG	A CG	CAAACTCCGACTACA	— Jβ1 2
AN6 2	Vβ3	—	TCGG CGG		— Jβ1 2
V11 5	Vβ3	— C	A		— Jβ1 2
C F6	Vβ3	—	A TG		— Jβ1 2
AN14 4	Vβ3	—	AACGC		— Jβ1 2
5C C7	Vβ3	—	A ATG		— Jβ1 2
AP15 2	Vβ1	—	TCGT G		— Jβ1 2
2B4	Vβ3	—	A T AG *	CC AGA AC C GT	— Jβ2 5
B10	VβB10	— C	CC CA	AACT TG T GC GT	— Jβ2 1
4 C3	VβB10	C	TTA CAGT	CT TG T GC GT	— Jβ2 1
V1 9 2	Vβ1	C A	A G	A A AGT T	— Jβ1 1
V15 4	Vβ8 3		GATGCG AA	TTT C GAAAGA TAT	— Jβ1 4

Fig 2. V_β junctional nucleotide sequences of T-cell receptors recognizing cytochrome *c* Sequences of the C F6, B10, 4 C3, 5C C7 T-cell clones[22] and 2B4 (refs 19, 20) and V1 9 2 (refs 19, 21) hybridomas have been reported previously and are included for comparison Only nucleotides in and around the diversity (*D*) gene segment are shown, *V* and *J* sequences both up- and downstream of this region were found to be identical to those published previously[5,19,21,49] *V*, *D* and *J* region boundaries were determined by comparison to known germline sequences ($V_\beta3$ (ref 19), $D_\beta1$ 1 (refs 31, 32), $J_\beta1$ 1, 1 2 and 1 4 (ref 5), $J_\beta2$ 1 and 2 5 (refs 5,19)) and to the cDNA sequences of BW5147 ($V_\beta1$, ref 21), 86T1 ($V_\beta1$, ref 49) and TB23 ($V_\beta8$ 3, ref 21) Dots, identity to the AP11 2 sequence or to the $D_\beta1$ 1 sequence The T nucleotide in the *D* segment of 2B4 (*) corresponds to the T at position 6 of $D_\beta2$ 1 (GGGAC*T*GGGGGGGC, refs 31, 32)

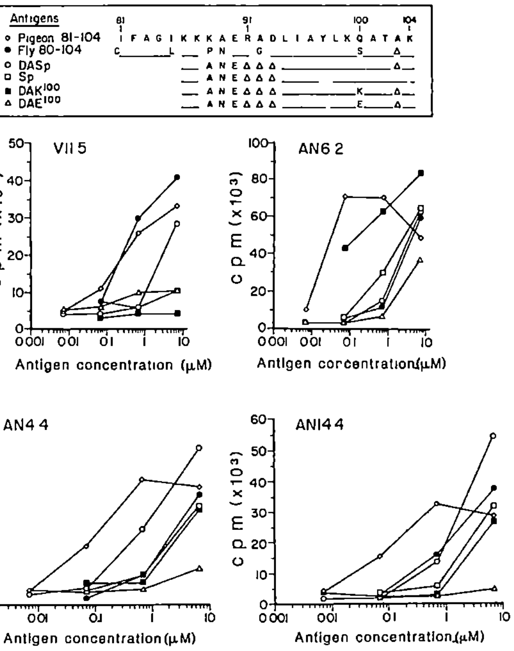
Methods The V_β genes of T-cell clones C F6, B10, 4 C3 and 5C C7 (ref 22) and hybridomas 2B4 (refs 19, 20) and V1 9 2 (refs 19, 21) were isolated from cDNA libraries as described The other V_β genes were isolated from limited genomic libraries constructed in the λL47 1, EMBL3, λgt7*lac5* or λ590 vectors and sequenced in the same manner as the V_α genes

during early development at the time of rearrangement of the T-cell receptor genes[27-30] In contrast, N-region diversification is never seen in immunoglobulin light chains[24,25], which also lack *D* gene segments, presumably because TdT is not present in these cells when the gene segments are rearranged TdT shows a clear preference for the addition of G residues[21] and three out of five of the asterisked sequences in Fig 1 contain G as their sole or predominant nucleotide Although no conclusions can be reached about the five sequences denoted in Fig 1 as '+' because we do not have the corresponding germline V_α and/or J_α gene segment sequences, N-region diversification may add up to 15 bases in the joining process[24,25] The two remaining unmarked sequences in Fig 1 can be totally accounted for on the basis of germline sequence and thus represent 'error-free' *V–J* joining This agrees with the suggestion[19] that the use of N-region diversification or the joining of a D_α gene segment is not required for productive recombination

It is striking that all the rearranged V_α and V_β genes are different in sequence except for V_α AN6 2 and V_α V11 5 This enormous variability arises through (1) use of a multiplicity of distinct gene segments, (2) joining these gene segments in different combinations and (3) diversity of junctional regions arising because of N-region diversification or the flexibility in the regions at which the gene segments are joined There is extensive junctional flexibility in the V_α and V_β genes and N-region diversification in the V_β genes[31,32] (Figs 1 and 2) This variability emphasizes the fact that striking diversity in the third hypervariable region is compatible with T-cell receptor specificity for particular antigen-MHC determinants

There is no simple correlation between class II MHC restriction and T-cell receptor gene segment usage (Table 1) For example, in helper T cells raised in $E^k_\alpha E^k_\beta$ mice, the most frequently studied haplotype, three different V_α (11 1, 11 2, and X), three different J_α (84, 14 4 and 2B4), five different V_β (3, 14, B10, 8 and Y) and three different J_β (1 2, 2 1 and 2 5) gene segments are used Moreover, the $V_\alpha11$ 1 gene segment is used by all four haplotypes studied, and the $V_\beta8$ and $V_\beta3$ gene segments are present in three of these four haplotypes The enormous diversity observed in the junctional regions of α and β chains suggests that there are many different ways to select the T-cell receptor molecule for restriction to a particular MHC molecule

Analysis of the ten $V_\alpha11$ 1 gene segments reveals no nucleotide differences (Fig 1 and data not shown) Since all ten T-helper hybridomas are functional and presumably represent mature T cells, we conclude that somatic hypermutation occurs infrequently, if at all, in the α genes of T_H cells specific for cytochrome *c* Similar conclusions have been reached for the β genes of the T-cell receptor[9 21,33] and our data are consistent with this interpretation in that eleven V_β gene segments from three subfamilies are identical in their coding region nucleotide sequences to their subfamily members, except for one mutation at position 95 of the V11 5 V_β gene segment (Fig 2) Somatic hypermutation in B cells, essential for affinity maturation[34], appears to occur relatively late in B-cell differentiation[35-37] Our T cells were hybridomas generated from purified lymph node cells that were stimulated once *in vivo* and a week later placed in tissue culture with antigen for three days before fusion Thus the possibility that somatic hypermutation has not occurred because our T cells are analogous to primary B cells or because



Antigens	81										91									100				104
○ Pigeon 81-104	I	F	A	G	I	K	K	K	A	E	R	A	D	L	I	A	Y	L	K	Q	A	T	A	K
● Fly 80-104	C	–	–	–	L	–	–	P	N	–	–	G	–	–	–	–	–	–	–	S	–	–	Δ	–
○ DASp						–	–	A	N	E	Δ	Δ	Δ	–	–	–	–	–	–	–	–	–	Δ	–
□ Sp						–	–	A	N	E	Δ	Δ	Δ	–	–	–	–	–	–	–	–	–	–	–
■ DAK¹⁰⁰						–	–	A	N	E	Δ	Δ	Δ	–	–	–	–	–	–	K	–	–	Δ	–
△ DAE¹⁰⁰						–	–	A	N	E	Δ	Δ	Δ	–	–	–	–	–	–	E	–	–	Δ	–

Fig 3 Fine specificity of cytochrome *c* peptide-specific T hybridomas Hybridoma cells were cultured with varying concentrations of pigeon cytochrome *c* and B10 A ($E^k_\alpha E^k_\beta$) spleen cells as a source of antigen presenting cells After 40 h of culture, supernatants were transferred to the IL-2-dependent cell line HT-2 ^{3}H-thymidine was added to the secondary culture 16 h later Antigens were C-terminal peptides of pigeon or fly cytochrome *c*, or synthetic derivatives of the pigeon peptide[46,50] Note that an N-terminal cysteine has been added to the fly sequence A dash denotes identity, Δ denotes a deletion (compared with the pigeon sequence)

of the antigen excess condition *in vitro* cannot be excluded However, the data here and data from 27 human V_β sequences from peripheral blood[38] are consistent with the supposition that somatic hypermutation does not play an important role in generating the diversity of the T-cell receptor repertoire Hence, affinity maturation may not play any significant role in the T-cell recognition of antigen-MHC and the affinity of T-cell receptors for their cognate antigen may be generally lower than that of antibodies for their cognate antigen

The V_α and V_β chains from hybridomas AN6 2 and V11 5 are identical except for five amino-acid differences in the V_β hypervariable region formed by *V-D-J* joining (Fig 2) We studied the antigenic fine specificity of these hybridomas to six different related cytochrome *c* peptides (Fig 3) These two T cells responded differently to the panel of peptides For example, hybridoma V11 5 did not respond to the peptide DAK^{100} even at high levels of antigen concentration, whereas hybridoma AN6 2 reacted strongly with this peptide Although no MHC fine specificity data is available for these two T-cell clones, we know that both respond equally well to antigen presented on cells bearing $\mathrm{E}^k_\alpha\mathrm{E}^k_\beta$ or $\mathrm{E}^k_\alpha\mathrm{E}^b_\beta$ MHC molecules and not to cells bearing $\mathrm{E}^k_\alpha\mathrm{E}^s_\beta$ or $\mathrm{E}^v_\alpha\mathrm{E}^v_\beta$ molecules We thus conclude that changes limited to the junctional region of the V_β sequence may alter the antigenic fine specificity of individual T cells without altering their predominant MHC specificity As this region seems to correspond to the third hypervariable segment of immunoglobulin V region sequences[21,33,39,40], we conclude that this region is very probably part of the antigen binding site—consistent with the supposition that the V regions of T-cell receptors and immunoglobulins fold in a similar manner.

Further fine specificity studies on two other hybridomas using six different cytochrome *c* peptides resulted in a striking observation Hybridomas AN4 4 and AN14 4, which do not share the same pair of V_α and V_β gene segments, show nearly identical profiles of antigen fine specificity within the panel of peptides tested (Fig 3, pigeon cytochrome *c* > DASp peptide > fly cytochrome *c* > Sp peptide > DAK^{100} peptide > DAE^{100} peptide) The fact that two hybridomas which are similar in their antigen fine specificity and predominantly restricted to $\mathrm{E}^k_\alpha\mathrm{E}^k_\beta$ do not share the same pair of V_α and V_β gene segments suggest that a combination of different V_α and V_β regions can form very similar antigen-MHC binding pockets, again consistent with the idea that affinity for cognate antigen may be relatively low

The V_α-restricted nature of the cytochrome *c* T-cell response raises interesting clinical possibilities Several animal models of human autoimmune diseases such as multiple sclerosis, rheumatoid arthritis and Hashimoto's thyroiditis provide strong evidence for the pathogenic involvement of T cells in the development of these autoimmune disorders[41-45] These models also indicate that T-cell responses responsible for these autoimmune diseases are directed against restricted antigen determinants These situations may be similar to the T-cell response against cytochrome *c* and raise the possibility that these autoimmune diseases may be associated with predominant usage of a particular T-cell receptor V_α (or V_β) gene segment Antibodies to these *V* gene segments might then have useful diagnostic or therapeutic possibilities

We thank Debra Maloney and Karyl Minard for technical assistance, Bertha Jones, Connie Katz and Cathy Elkins for preparation of the manuscript, and Drs Richard Barth, Patrick Concannon, Joan Kobori, Mitchell Kronenberg, Ulf Landegren and Nilabh Shastri for a critical reading of the manuscript This research was supported by the NIH (L H) and NIH grant-in-aid AI-20661 and CA-06927 (D H)

Received 18 July, accepted 17 October 1986

1 Kindred, B & Sheffler, D C *J Immun* **109**, 940-943 (1972)
2 Katz, D H , Hamaoka, T & Benacerraf, B *J exp Med* **137**, 1405-1418 (1973)
3 Rosenthal, A S & Shevach, E M *J exp Med* **138**, 1194-1212 (1973)
4 Zinkernagel R M & Doherty P C *Nature* **248**, 701-702 (1974)
5 Gascoigne, N R J , Chien, Y-H , Becker, D M , Kavaler, J & Davis, M M *Nature* **310**, 387-391 (1984)
6 Malissen, M *et al. Nature* **311**, 344-350 (1984)
7 Siu, G *et al. Cell* **37**, 393-401 (1984)
8 Toyonaga, B , Yoshikai, Y , Vadasz, V , Chin, B & Mak, T W *Proc. natn Acad Sci U S A* **82**, 8624-8628 (1986)
9 Chien, Y-H *et al. Nature* **312**, 31-36 (1984)
10 Saito, H *et al. Nature* **312**, 36-40 (1984)
11 Sim, G K *et al. Nature* **312**, 771-775 (1984)
12 Heber-Katz, E *et al. J exp Med* **155**, 1086-1099 (1982)
13 Mengle Gaw, L , Conner, S , McDevitt, H O & Fathman, C G *J exp Med* **160**, 1184-1194 (1984)
14 Mengle-Gaw, L & McDevitt, H O *Proc. natn Acad Sci U S A.* **82**, 2910-2914 (1985)
15 Schwartz, R H *A. Rev Immun* **3**, 237-261 (1985)
16 Solinger, A M , Ultee, M E , Margoliash, E & Schwartz, R. H *J exp Med* **150**, 830-848 (1979)
17 Hansburg, D , Fairwell, T , Schwartz, R. H & Appella, E *J Immun* **131**, 319-324 (1983)
18 Hansburg, D *et al. J Immun* **127**, 1844-1851 (1981)
19 Chien, Y H , Gascoigne, N J , Kavaler, J , Lee, N E & Davis, M M *Nature* **309**, 322-326 (1984)
20 Becker, D *et al. Nature* **317**, 430-434 (1985)
21 Barth, R *et al. Nature* **316**, 517-523 (1985)
22 Fink, P J , Matis, L A , McElligott, D L , Bookman, M & Hedrick, S M *Nature* **321**, 219-226 (1986)
23 Arden, B , Klotz, J L , Siu, G & Hood, L *Nature* **316**, 783-787 (1985)
24 Sakano, H , Maki, R , Kurosawa, Y , Roeder, W & Tonegawa, S *Nature* **286**, 676-683 (1980)
25 Alt, F & Baltimore, D *Proc. natn Acad Sci. U S A* **78**, 5812-5816 (1982)
26 Alt, F W & Baltimore, D *Proc. natn Acad Sci U S A* **79**, 4118-4122 (1982)
27 Chang, L M S *Biochem biophys Res Commun.* **44**, 124-131 (1971)
28 Kung, P C , Silverstone, A E , McCaffrey, R P & Baltimore, D *J exp Med* **141**, 855-865 (1975)
29 McCaffrey, R P , Harrison, T A , Parkman, R & Baltimore, D *N Engl J Med* **292**, 775-780 (1975)
30 Rothenberg, E & Triglia, D *J Immun.* **130**, 1627-1633 (1983)
31 Kavaler, J , Davis, M M & Chien, Y-H *Nature* **310**, 421-423 (1984)
32 Siu, G *et al. Nature* **311**, 344-350 (1984)
33 Behlke, M A *et al Science* **229**, 566-570 (1985)
34 Bauer, D C , Mathies, M J & Stravitsky, A A *J exp Med* **117**, 889-907 (1963)
35 Weigert, M G , Cesari, I M , Yonkovich, S J & Cohn M *Nature* **228**, 1045-1047 (1970)
36 Kim, S , Davis, M , Sinn, E , Patten, P & Hood, L *Cell* **27**, 573-581 (1981)
37 Pech, M , Hochtl, J , Schnell, H & Zachau, H G *Nature* **291**, 668-670 (1981)
38 Concannon, P , Pickering, L A , Kung, P & Hood, L *Proc. natn Acad Sci U S A.* **83**, 6598-6602 (1986)
39 Patten, P *et al. Nature* **312**, 40-46 (1984)
40 Wu, T & Kabat, E *J exp Med* **132**, 211-250 (1970)
41 Zamvil, S *et al. Nature* **317**, 355-358 (1985)
42 Ber-nun, A & Cohen, I R *J Immun* **129**, 303-308 (1982)
43 Holoshitz, J , Napoarstek, Y , Ben-nun, A & Cohen, I R *Science* **219**, 56-58 (1982)
44 Maron, R , Zerubavel, R , Friedman, A & Cohen, I R *J Immun* **131**, 2316-2322 (1983)
45 Holmdahl, R , Klareskog, L , Rubin, K., Larsson, E & Wigzell, H *Scand. J Immunol* (in the press)
46 Hansburg, D & Appella, E *J Immun* **135**, 3712-3718 (1985)
47 Winoto, A , Mjolsness, S & Hood, L *Nature* **316**, 832-836 (1985)
48 Strauss, E C , Kobori, J A , Siu, G & Hood, L *Analyt Biochem.* **154**, 353-360 (1986)
49 Hedrick, S M , Nielsen, E A , Kavaler, J , Cohen, D I & Davis, M M *Nature* **308**, 153-158 (1984)
50 Hansburg, D , Heber-Katz, E , Fairwell, T & Appella, E *J exp Med* **158**, 25-39 (1983)

High frequency of unequal recombination in pseudoautosomal region shown by proviral insertion in transgenic mouse

Klaus Harbers*, Philippe Soriano†, Ursula Muller* & Rudolf Jaenisch†

* Heinrich-Pette-Institut fur Experimentelle Virologie und
Immunologie an der Universitat Hamburg, Martinistrasse 52,
2000 Hamburg 20, FRG
† Whitehead Institute for Biomedical Research and Department of
Biology, Massachusetts Institute of Technology,
Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

The mammalian X and Y chromosomes, in contrast to the autosomes, pair during male meiosis only near the telomeres. Alleles localized in this region can undergo reciprocal exchange during meiosis[1-4]. Because such sequences do not show strict sex-linked inheritance, they have been termed pseudoautosomal[5] In man, several DNA sequences have been described which show pseudoautosomal transmission and which are localized in the pairing region at the ends of the short arms of both the X and Y chromosomes (refs 6-9, and D. Page, unpublished results). We now show that the transgenic mouse strain, Mov-15, contains a single Moloney murine leukaemia virus (M-MuLV) genome in its

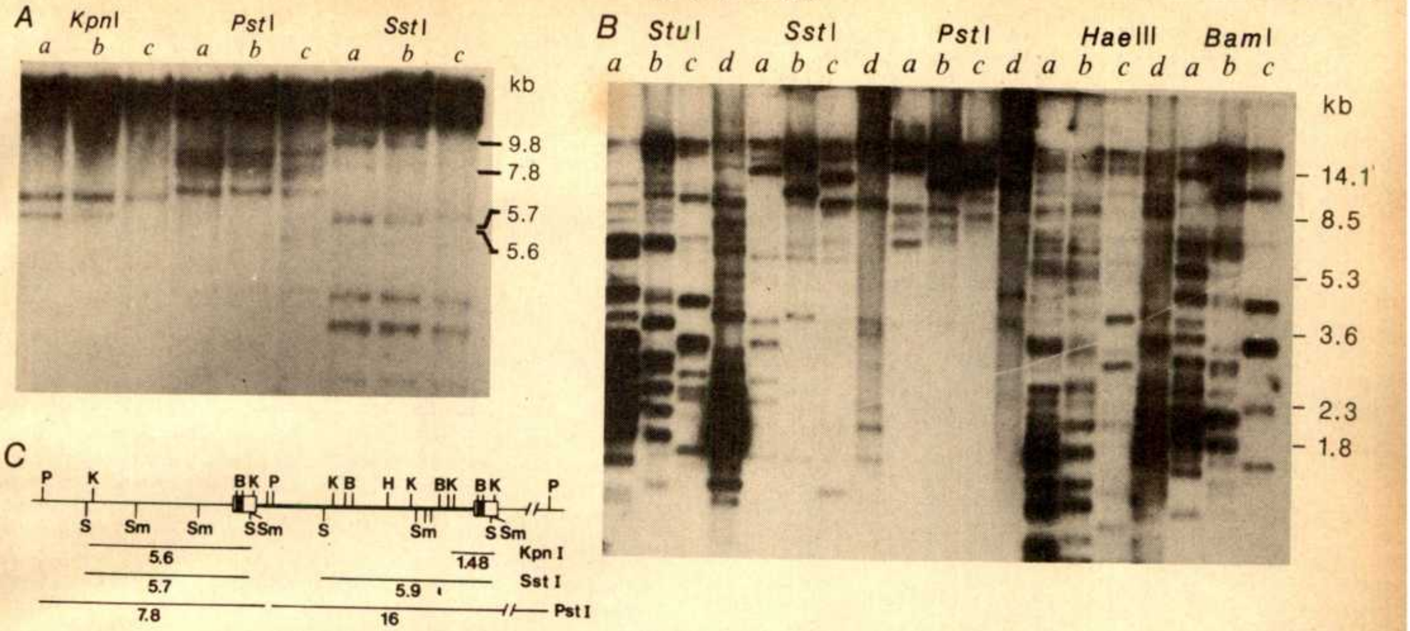


Fig. 1 Restriction map and cloning of the Mov-15 locus. *A*, Isolation of mouse DNA sequences flanking the Mov-15 provirus. A recombinant plasmid, pMov-15/1, containing sequences flanking the provirus, was hybridized to DNAs from homozygous Mov-15 males (lane *a*), heterozygous Mov-15 males (lane *b*) and BALB/c males (lane *c*) digested with the indicated restriction enzymes. *B*, DNA polymorphism at the Mov-15 locus. DNAs from BALB/c (lanes *a*), C57BL/6 (lanes *b*), 129/J (lanes *c*), and SWR/J (lanes *d*) mice were digested with the restriction enzymes shown and analysed by Southern blot analysis as described in *A*, using pMov-15/1 as a probe. *C*, Restriction map of the proviral region. White boxes, LTRs; black boxes, the bacterial *supF* gene; black bar, the proviral genome and the line left and right from the provirus, the flanking mouse sequences. K, P, H, S and Sm denote cleavage sites for *Kpn*I, *Pst*I, *Hin*d III, *Sst*I and *Sma*I, respectively. The map was deduced from restriction enzyme analysis of cellular DNA from homozygous Mov-15 mice. The lengths of internal and flanking fragments hybridizing with the bacterial suppressor gene are indicated in kb below the figure.

Methods. DNAs were purified, digested with the enzyme indicated, separated on a 0.8% gel and transferred to Gene Screen Plus membranes (NEN)[10]. The filters were hybridized to probes labelled to high specific activity by random oligonucleotide priming[26]. Filters were washed in 0.1 × SSC, 0.1% SDS, at 68 °C. As predicted from the map in *C*, hybridization of the pMov-15/1 probe revealed the presence of a 5.6 kb *Kpn*I fragment, a 7.8 kb *Pst*I fragment and a 5.7 kb *Sst*I fragment in Mov-15 mice (lanes *a* and *b*). These fragments were not seen in DNAs from wild-type mice (lane *c*) and the results therefore confirm that pMov-15/1 represents cellular sequences flanking the provirus. The 9.8 kb *Sst*I fragment found exclusively in DNAs from Mov-15 mice (lanes *a* and *b*) most probably represents a flanking cellular fragment from the 3′ end of the provirus which hybridized with pMov-15/1 due to the repetitive nature of the probe. As expected, the intensities of the Mov-15 specific fragments in homozygous animals were about twice that seen in DNA from heterozygous animals (compare lanes *a* and *b*). The 5.7 kb *Sst*I fragment had been cloned by extracting DNA from the liver of a homozygous Mov-15 female mouse, digesting it with *Sst*I and fractionating it in a sucrose gradient. Fractions containing the Mov-15 specific *Sst*I fragments were identified by hybridization with the *supF* gene and ligated to *Sst*I-cleaved λ gtWES arms. After *in vitro* packaging, recombinant phage particles were titred on the sup^+ host LE 392. About 1.5×10^6 plaque-forming units were then plated on the sup^- host MC 1061 p3[12]. Eight plaques were obtained, five of which contained recombinant phages with the 5.7 kb flanking *Sst*I fragment. The remaining three plaques contained recombinant phages with the internal 5.9 kb *Sst*I fragment. This internal fragment was distinguished from the flanking *Sst*I fragment on the basis of an internal *Hin*dIII site at nucleotide 4894 (ref. 27). A recombinant λ phage containing the flanking sequences was cleaved with *Sst*I, the 5.7 kb insert isolated by preparative gel electrophoresis, further digested with *Bam*HI to remove most of the LTR, and the *Bam*HI–*Sst*I fragment comprising about 90 bp of the LTR plus the flanking mouse sequences was subcloned in pUC 19. The resulting plasmid was designated pMov-15/1.

germline, and genetic evidence indicates that the provirus is integrated into the pseudoautosomal region of the sex chromosome. Proviral copies are lost or gained in 7% of male meioses in this strain, and mouse sequences flanking the provirus are tandemly repeated and highly variable. We conclude that unequal recombination events occur with high frequency in the pairing region, possibly because of the presence of repeated sequences.

The derivation of the Mov-15 strain from $\text{M-MuLV}^{\text{sup}}$-infected preimplantation BALB/c mouse embryos has been described previously[10]. Since all Mov-15 animals develop viraemia, the provirus serves as a dominant marker whose expression can be monitored by radioimmunoassay of serum samples[11]. Animals carrying the provirus were also identified by DNA blot analysis, using as a probe the bacterial suppressor gene present in the long terminal repeats (LTRs)[12]. When two Mov-15 males of the N1 generation (obtained from breeding founder male A, ref. 10) were bred with wild-type BALB/c females, about 90% of male offspring and only 10% of female offspring carried the Mov-15 provirus (Table 1). This partial sex linkage suggested that the provirus in these males was located on the Y chromosome and was transferred to the X chromosome in about 10% of the meioses, resulting in 10% viraemic daughters. To test this hypothesis, viraemic N2 females were bred with wild-type BALB/c males. As shown in Table 1, the Mov-15 provirus was transmitted equally to sons and daughters of the N3 generation. When viraemic males of the N3 generation which had received the provirus from their mother were bred with wild-type BALB/c females, partial X-linked transmission of the virus was observed (85% of the N4 daughters and only 19% of the N4 sons carried the virus), as expected if the provirus was carried on the X chromosome of the N2 mothers. Viraemic sons of the N4 generation in turn transmitted the virus in a Y-linked mode to the next generation. The results clearly show that the Mov-15 provirus is transferred from the Y to the X chromosome and back from the X to the Y chromosome in 10–20% of all male meioses (Table 1). A gradient of sex linkage has been established within the human pseudoautosomal region, and the recombination frequency can be used to estimate the position of the locus within this region[9]. Thus, markers at the telomeres recombine with frequencies of 50%, and the recombination rate decreases for more proximal markers. If the same holds true for the mouse pseudoautosomal region, our data localize the Mov-15 proviral

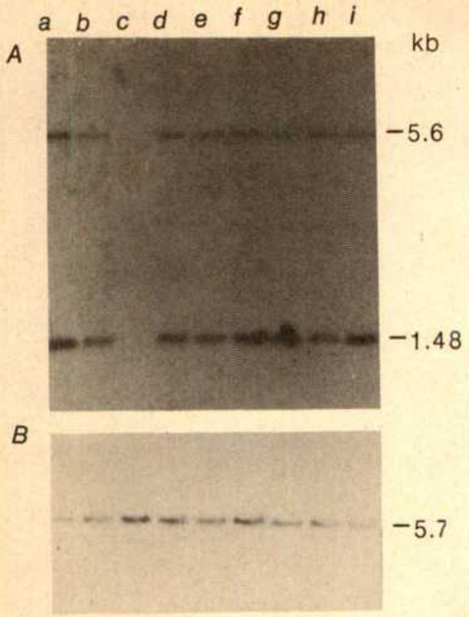


Fig. 2 Unequal recombination at the Mov-15 locus. A homozygous Mov-15 male was mated with a wild-type BALB/c female and embryos were isolated at day 14 of development. DNAs were purified and analysed by blot hybridization following digestion with *Kpn*I (Fig. 1 legend). *A*, Hybridization to πAN7, a plasmid containing the bacterial suppressor gene (gift of H. Huang). The flanking 5.6 kb fragment and the internal 1.48 kb fragment hybridizing to the probe (Fig. 1*C*) are indicated. Since viraemia might lead to the presence of additional proviral copies in the DNA samples, which would all contain the internal fragment, the intensity of the 5.6 kb flanking fragment, which is specific for the provirus inserted in the germ line, was used instead to estimate the number of Mov-15 proviral copies present. *B*, Hybridization with a *Bgl*II–*Hpa*I fragment containing part of the $\alpha 1$(I) collagen gene[28]; the intensity of the 5.7 kb hybridizing fragment serves as a control for the amount of DNA loaded in each lane.

Table 1 Sex-chromosome-linked transmission of Mov-15 provirus

Parents		Offspring			
Inferred genotype F	M	Backcross generation	Sex	Offspring with provirus/total tested	%
wt/wt	Y^{Mov}/X^{wt} (N1)	N2	M	116/130	89
			F	19/191	10
X^{Mov}/X^{wt} (N2)	wt/wt	N3	M	31/57	54
			F	33/51	64
wt/wt	Y^{wt}/X^{Mov} (N3)	N4	M	19/99	19
			F	130/152	86
wt/wt	Y^{Mov}/X^{wt} (N4)	N5	M	43/49	88
			F	14/70	20

All offspring from crosses of Mov-15 males with wild-type BALB/c females were analysed for the presence of virus in the serum by radioimmunoassay using antibodies against p30 (ref. 11). In addition, about 10% of the offspring were analyzed for the presence of proviral sequences by DNA blot hybridization. No discrepancies were observed when the results of both tests were compared. Since virus can be efficiently transmitted by milk from a viraemic mother to her offspring[24], all offspring of Mov-15 females with wild-type BALB/c males (N3) were analysed for the presence of the provirus by blot analysis of tail DNA samples. Among the offspring of the original founder animal (founder A, ref. 10), 3 out of 19 male offspring and 7 out of 27 female offspring contained the Mov-15 provirus. Because of the relatively small number of provirus-carrying offspring, a clear assignment of the provirus in the founder animal to the X or to the Y chromosome could not be made, but two of the three provirus-carrying males were found to have the provirus on the Y chromosome (lines 1 and 2). wt, wild-type; M, males; F, females.

genome to the proximal part of the X–Y pairing region. Three point crosses with distal markers such as the steroid sulphatase gene[3] will allow us to test this hypothesis directly.

In order to analyse the organization of the Mov-15 locus, a probe corresponding to sequences flanking the provirus was isolated by cloning into λgt WES using selection for the *supF* gene (ref. 12; see Fig. 1*A* legend). This probe, designated pMov-15/1, hybridized to differently-sized fragments in virus-positive and virus-negative animals, demonstrating that the cloned sequences represent cellular sequences flanking the Mov-15 provirus (Fig. 1*A*). In addition to the Mov-15 specific fragments, the probe hybridized with variable intensities to many other fragments of different sizes, even under stringent conditions of hybridization and washing. To analyse the flanking sequences further, the insert of pMov-15/1 was digested with *Sma*I and the three fragments produced (Fig. 1*C*) were sub-cloned and used as probes on Southern blots. Each fragment produced identical patterns of hybridization to those shown in Fig. 1*A*. Each fragment hybridized with all the others, confirming that the cellular sequences flanking the provirus are not only repeated many times in the genome, but are also internally repetitive. The repeat length unit is at the most 1.3 kilobases (kb), the length of the smallest fragment tested.

The flanking sequences were also used to probe DNAs from different mouse strains. Hybridization bands of different size and intensity were observed when DNAs from different mouse strains were compared (Fig. 1*B*). This polymorphism was not detected when males and females of the same inbred strain (BALB/c) were compared (data not shown). The sequences are specific for mouse and do not hybridize, even under low stringency, to DNAs from man, hamster, rat, pig and chicken (data not shown). Repetitive and highly polymorphic sequences have been found previously in the human pseudoautosomal region (refs 6–9, and D. Page, unpublished results) as well as on the mouse Y chromosome[13,14].

To derive animals homozygous for virus integration at the Mov-15 locus, heterozygous males carrying the provirus on their Y chromosome were crossed with heterozygous females. The genotypes of offspring were determined by Southern blot analysis using the autoradiographic intensity of the 5.6 kb Mov-15-specific *Kpn*I fragment (representing part of the LTR and 5′ flanking sequences; Fig. 1*C*) to estimate the number of proviral copies present. Animals with two proviruses were classified as homozygous. As predicted, about equal numbers of homozygous and heterozygous sons and of heterozygous and wild-type daughters resulted from this cross; no wild-type sons or homozygous daughters were found (Table 2).

Table 2 M-MuLV genotype of offspring from parents both heterozygous at the Mov-15 locus

	Cross: $F(X^{wt}/X^{Mov}) \times M(X^{wt}/Y^{Mov})$*		
	Genotype of offspring		
	Homozygous	Heterozygous	Wild type
M	18	22	0
F	0	14	18

The genotype of the offspring was determined by Southern blot analysis of tail DNA using the autoradiographic intensity of the 5.6 kb Mov-15 specific *Kpn*I fragment for estimating the number of M-MuLV^{sup} copies present (for details see legend to Fig. 2). wt, wild-type; M, males; F, females.

* The genotype X^{wt}/Y^{Mov} was assigned to the heterozygous males since they transmitted the Mov-15 provirus to about 90% of their male offspring and about 10% of their female offspring when mated with wild-type BALB/c females.

When males homozygous for virus integration were mated with wild-type BALB/c females, only 93% of the offspring were viraemic instead of the 100% expected for normal Mendelian transmission (Table 3). Southern blot analysis using the *supF* gene fragment or a representative M-MuLV as a probe confirmed that non-viraemic animals did not carry the provirus. As an example, Fig. 2 shows the analysis of embryos derived from one such cross. One of the embryos was negative for proviral sequences (lane *c*). No other fragments hybridizing with the *supF* probe were detected, suggesting that the absence of the provirus was not due to homologous recombination involving the LTRs, as such a mechanism of provirus excision should leave a single LTR at the integration site[15]. While the majority of the offspring carried a single Mov-15 provirus, some animals had acquired two copies. This was shown by comparing the intensities of the 5.6 kb *Kpn*I fragments with a unique fragment as an internal standard (see lane *a* in Fig. 2*A*,*B*). Similar analyses of a larger number of embryos showed that about 7% had acquired two proviral copies (Table 3). The frequency of offspring which have lost or gained a proviral copy is therefore almost as high as the overall frequency of recombination between the X and Y chromosomes at the Mov-15 locus.

The results in Table 3 and Fig. 2 suggest that the region of the X and Y chromosomes into which the provirus has integrated can undergo unequal exchange during male meiosis, as indicated by the loss of the provirus on one sex chromosome (provirus negative offspring) and the gain of a second copy on the other sex chromosome (offspring with two proviral genomes). Blot hybridization analyses on pulse field gels demonstrated that restriction sites in cellular sequences flanking the provirus were identical over a distance of 70 kb (data not shown). Thus the recombination events involved large stretches of host sequences and the provirus was not reinserted into an unrelated part of the genome.

The most likely mechanism to explain the gain or the loss of proviral genomes is unequal crossing-over events between sister chromatids or between homologues either during gametogenesis or during meiosis. Such a mechanism would indeed predict that the two classes of offspring, either with no provirus or with two proviral copies, should arise at equal frequencies. To date, no direct genetic information on the occurrence and frequency of unequal crossing-over is available for mammals. The only mechanism known to occur in mammals which formally could result in unequal crossing-over is sister chromatid exchange. The frequency of sister chromatid exchange in spermatogonia has been determined to be less than two per cell for any chromosomes[16], which is therefore orders of magnitude lower than the frequency we have observed for the unequal exchange of the Mov-15 provirus which occurs within a small chromosomal region. The most likely mechanism by which the Mov-15 provirus is deleted or duplicated is therefore unequal crossing-over during meiosis. Because each recombination event at the Mov-15 locus transfers at least 70 kb of sequences flanking the provirus from one sex chromosome to the other, our observations suggest that a high frequency of unequal recombination may be characteristic for the X-Y chromosome pairing region. In addition, the recombination frequencies we have determined in this work have been in male meiosis only. In the human pseudoautosomal region it has been estimated that the rate of recombination is tenfold higher in males than in females[9]. It will therefore be of interest to study recombination as well as unequal crossover frequencies in Mov-15 females.

Table 3 Transmission of Mov-15 provirus

Parents		Offspring carrying:	
		no provirus	two proviruses
F (wt/wt) x M(X^{Mov}/Y^{Mov})	M:	3/76	
	F:	10/115	
	E:	7/109	8/109
	Total:	20/300 (6.7%)	8/109 (7.3%)

Males homozygous for virus integration at the Mov-15 locus were derived as described in Table 2 and mated to wild-type females. Offspring were tested for viraemia at the age of 4–6 weeks. Non-viraemic animals were analysed for the presence of the provirus by blot hybridization of tail DNAs. In addition, the same tail DNAs were analysed for the presence of the M-MuLV specific 5.9 kb and 2.6 kb *Sst*I fragments as described[25]. Embryo DNAs (isolated between day 12 and 14 of gestation) were analysed for the presence of the provirus by blot hybridization and the intensity of the Mov-15 specific 5.6 kb *Kpn*I fragment on autoradiograms was used for estimating the number of M-MuLV copies present (for details see legend to Fig. 2). wt, wild-type; M, males; F, females; E, embryos.

It has been suggested that repetitive sequences can lead to unequal crossover events resulting in contraction or expansion of sequence families[17]. If such a mechanism is responsible for the unequal exchange observed at the Mov-15 locus, it may be possible to generate animals with larger numbers of proviral copies by successive breeding of offspring with newly-acquired Mov-15 proviruses and so allow comparison of unequal crossing-over frequencies in animals after 'expansion' or 'contraction' of the Mov-15 pairing region. To date, unequal recombination has been directly documented in yeast[18-20] and *Drosophila*[21], but never in mammals, although it has been implicated in the evolution of multigene families[22] and satellite DNAs[17,23] and in the generation of autosomal[29] or sex-linked mutant loci[1]. No information has been available until now on the frequency of such events during meiosis. Only the presence of the unique proviral marker has enabled us to measure the frequency of unequal recombination in a DNA region composed of repetitive sequences. Our results suggest that such exchanges may be frequent in regions where chromosomes synapse during meiosis.

We thank Kirsten Haase and Doris Grotkopp for technical assistance, our colleagues for critical comments on the manuscript, and David Page and Tom Gridley for many helpful suggestions. Supported by grants from the Deutsche Forschungsgemeinschaft, the Stiftung Volkswagenwerk, grant HD-19105 from the National Institutes of Health, and by grant PO1-CA38497 from the National Cancer Institute.

Received 19 August; accepted 10 October 1986.

1. McLaren, A., Simpson, E., Tomonari, K., Chandler, P. & Hogg, H. *Nature* **312,** 552–555 (1984).
2. Simpson, E. *Cell* **44,** 813–814 (1986).
3. Keitges, E., Rivest, M., Siniscalco, M. & Gartler, S. M. *Nature* **315,** 226–227 (1985).
4. Polani, P. E. *Hum. Genet.* **60,** 207–211 (1982).
5. Burgoyne, P. S. *Hum. Genet.* **61,** 85–90 (1982).
6. Cooke, H. J., Brown, W. R. A. & Rappold, G. A. *Nature* **317,** 687–692 (1985).
7. Simmler, M. C. *et al. Nature* **317,** 692–697 (1985).
8. Buckle, V., Mondello, C., Darling, S., Craig, I. W. & Goodfellow, P. N. *Nature* **317,** 739–741 (1985).
9. Rouyer, F. *et al. Nature* **319,** 291–295 (1986).
10. Soriano, P. & Jaenisch, R. *Cell* **46,** 19–29 (1986).
11. Strand, M., August, J. T. & Jaenisch, R. *Virology* **76,** 886–890 (1977).
12. Reik, W., Weiher, H., Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* **82,** 1141–1145 (1985).
13. Phillips, S. J., Birkenmeier, E. H., Callahan, R. & Eicher, E. M. *Nature* **297,** 241–243 (1982).
14. Lamar, E. E. & Palmer, E. *Cell* **37,** 171–177 (1984).
15. Copeland, N. G., Hutchison, K. W. & Jenkins, N. A. *Cell* **33,** 379–387 (1983).
16. Allan, J. W. & Latt, S. A. *Nature* **260,** 449–451 (1976).
17. Smith, G. P. *Science* **191,** 528–535 (1976).
18. Petes, T. D. *Cell* **19,** 765–774 (1980).
19. Szostak, J. W. & Wu R. *Nature* **284,** 426–430 (1980).
20. Jackson, J. A. & Fink, G. R. *Genetics* **109,** 303–332 (1985).
21. Tartof, K. D. *Cold Spring Harb. Symp. quant. Biol.* **38,** 491–500 (1974).
22. Baltimore, D. *Cell* **24,** 592–594 (1981).
23. Jeffreys, A. J., Wilson, V. & Thein, S. L. *Nature* **314,** 67–73 (1985).
24. Jaenisch, R. *Cell* **19,** 181–188 (1980).
25. Jaenisch, R. *et al. Cell* **24,** 519–529 (1981).
26. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132,** 6–13 (1983).
27. Shinnick, T. M., Lerner, R. A. & Sutcliffe, J. G. *Nature* **293,** 543–548 (1981).
28. Harbers, K., Kuehn, M., Delius, H. & Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* **81,** 1504–1508 (1984).
29. Lyon, M., Evans, E., Jarvis, S. & Sayers, I. *Nature* **279,** 38–42 (1979).

In vivo protein–DNA interactions in a glucocorticoid response element require the presence of the hormone

Peter B. Becker, Bernd Gloss, Wolfgang Schmid, Uwe Strähle & Günther Schütz

Institute of Cell and Tumor Biology, German Cancer Research Center, Im Neuenheimer Feld 280, D-6900 Heidelberg, FRG

Transcriptional activation of gene expression by glucocorticoid hormones is mediated by the interaction of hormone–receptor complexes with specific DNA sequences called glucocorticoid responsive elements (GREs) (refs 1–3, see ref. 4 for review). Deletion of this sequence abolishes glucocorticoid induction of transcription[4-8]. According to a current model, activation of the cytoplasmic receptor protein by hormone binding leads to its increased affinity for and translocation to the nucleus[4]. However, recent reports that the oestradiol and progesterone receptors are localized in the nucleus in the absence of steroid[9-11] led us to examine whether the free receptor interacts *in vivo* with its DNA binding site in the absence of hormone binding. We used the genomic footprinting technique[12-15] to show that changes in *in vivo* protein–DNA interactions within the GREs of the tyrosine aminotransferase gene (*TAT*) can be detected only after hormone treatment in hepatoma cells. Such changes are not detected in fibroblast cells, in which the *TAT* gene is not expressed. Many of the changes in dimethylsulphate reactivity observed in the living cell are also found *in vitro* using cloned DNA and a partially purified glucocorticoid receptor.

The GREs of the rat *TAT* gene are located unusually far upstream, at −2.5 kilobases (kb), from the transcriptional start site (Fig. 1). These elements were initially indicated by the appearance of a DNaseI-hypersensitive site following hormone treatment[16] and then identified by gene transfer experiments and DNaseI footprinting using partially purified glucocorticoid receptor complex *in vitro*[17]. Sequence analysis revealed several sequence motifs that fit the established GRE consensus sequence[1,2,4]. In this study we focus on those elements (II and III, Fig. 1) that have been shown to be essential for glucocorticoid induction by gene transfer experiments.

The instability during purification of the glucocorticoid receptor in the absence of ligand has hindered the analysis of its potential interaction with DNA sequences by *in vitro* methods, but the recent development of the genomic footprinting technique[12,15] has made possible the analysis of receptor-DNA interactions *in vivo* in the absence of steroid. Cells of the *TAT*-expressing hepatoma cell line FTO-2B were withdrawn from glucocorticoids by incubation for at least 16 hours in serum-free medium before induction with the steroid dexamethasone for various times. Trypsinized cells suspended in medium were reacted with dimethyl sulphate (DMS) to probe for the reactivity of the N7 position of guanines in the major groove of the DNA helix[15,19,20]. Patterns of guanine (G) residues were obtained using the probe fragment indicated in Fig. 1.

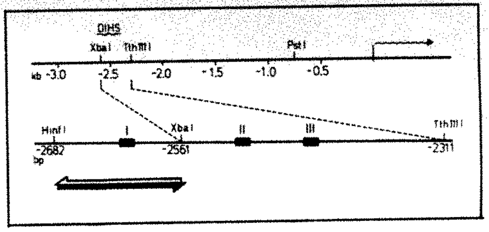


Fig. 1 The 5′-flanking region of the rat *TAT* gene. Upper line, 3 kb of the 5′-flanking region of the rat *TAT* gene: Numbering, distance from the cap-site; wavy arrow, the start of transcription; DIHS, region of dexamethasone-induced DNaseI hypersensitivity[16]. The lower line, enlargement of the region around the dexamethasone-inducible DNaseI-hypersensitive site: black boxes, sequence homologies to the GRE consensus sequence labelled with roman numerals; arrows, single-stranded DNA probes used to generate genomic sequence synthesized from an M13 vector containing the *Hin*fI-*Xba*I fragment.

Figure 2*a* shows an example of such an analysis of uninduced and induced FTO cells. The band intensity reflects the reactivity of a given guanine or a group of unresolved guanines for the chemical modification. Protection from methylation at guanines is most readily explained by protein binding whereas enhanced modification of guanines has been interpreted as a local increase in reagent concentration in a hydrophobic pocket caused by close contacts of protein to DNA[20]. Methylation protection and enhancement after hormone treatment were analysed quantitatively for each guanine (Fig. 2*c*). The data were collected from four independent experiments, each DNA sample being analysed on at least two blots. A number of changes in band intensity appear to be correlated with hormone treatment. DMS protections are evident at the sites of glucocorticoid interaction as identified by DNaseI footprinting[17]. Additional changes in reactivity map to neighbouring sequences (see Fig. 4). In a non-expressing rat fibroblast cell-line (XC cells) no such changes could be discerned after hormone treatment (Fig. 2*a*), nor were they observed when FTO-2B cells were treated with the glucocorticoid antagonist RU486 (Fig. 2*b*).

To ascertain whether the changes in DMS reactivity observed *in vivo* at the *TAT* gene GREs are caused by the glucocorticoid receptor, we performed *in vitro* methylation experiments, using cloned *TAT* DNA fragments containing the GREs and a partially purified glucocorticoid receptor (Fig. 3). The most prominent G protections map to the conserved TGTTCT motif and are identical with those shown to be contacted by the receptor in the GRE of mouse mammary tumour virus (MMTV)[21].

Figure 4 summarizes the data obtained *in vivo* and *in vitro*. Comparison of *in vivo* G reactivity patterns from uninduced and induced FTO-2B cells clearly shows that protection of guanines in the binding site is seen after dexamethasone treatment of cells (Fig. 2). They are accompanied by a number of hormone-induced changes in reactivity in neighbouring sequences. The most dramatically enhanced G reactivities are seen at four adjacent guanines (−2,459 to −2,456) between the two receptor binding sites. Most of the changes in DMS reactivity following hormone treatment are also found *in vitro* using a partially purified glucocorticoid receptor. The changes are not found in XC cells, where the *TAT* gene is not transcribed. These findings strongly indicate that changes in reactivity at guanines within the GRE are indeed caused by interaction of the glucocorticoid receptor with its binding site and that its affinity in the absence of hormone binding is too low to lead to a tight interaction with a GRE at receptor concentrations found in the living cell.

Receptor binding is accompanied by changes in chromatin structure that render a region of 200 base pairs (bp) surrounding the GREs hypersensitive to DNaseI digestion[16]. Whether this altered structure is a prerequisite for receptor binding or rather a consequence of binding is not known. For the MMTV GREs a similar DNaseI hypersensitivity has been shown to precede or parallel hormonal stimulation of MMTV promoter function[18].

It should be emphasized that these effects on guanine reactivities have been observed within a region which is unusually remote (at −2.5 kb) from the cap-site that is selectively rendered hypersensitive to DNase I digestion after dexamethasone treatment. Using the same DNA samples no hormone-dependent DMS effects were seen within several hundred nucleotides surrounding the *TAT* gene cap-site or within a region of DNaseI hypersensitivity at −1,000 bp (data not shown). Within the first

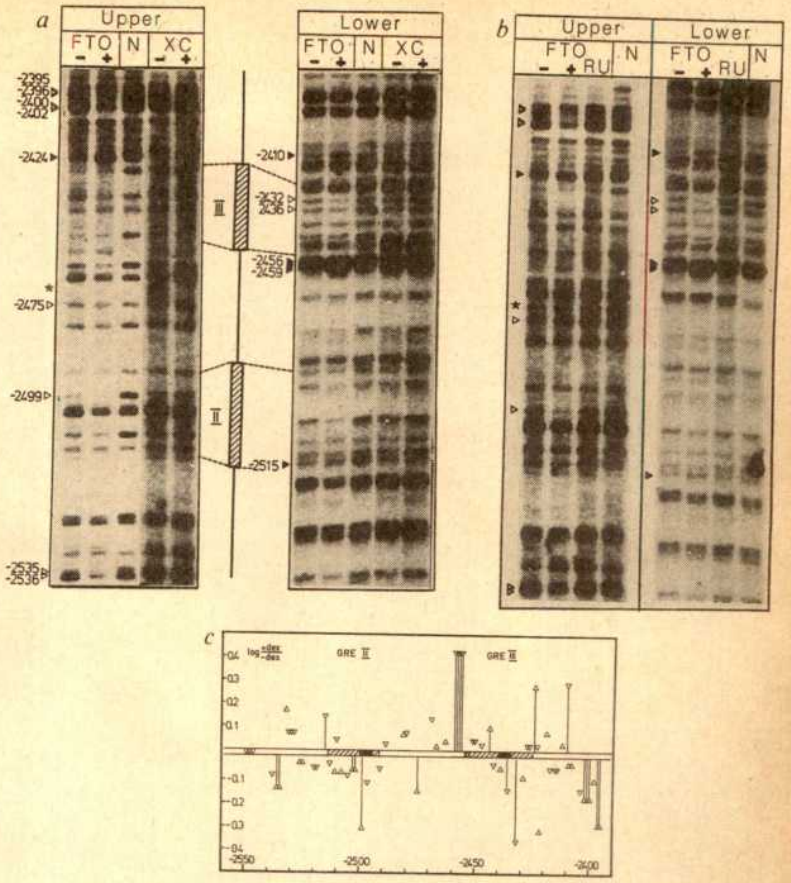
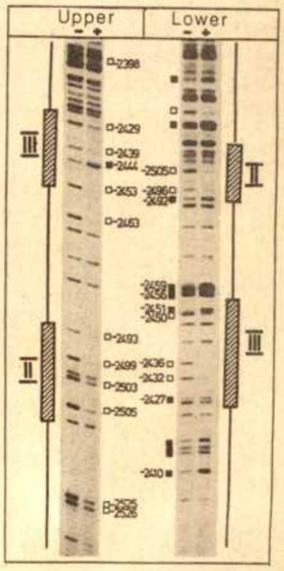


Fig. 2 Glucocorticoid-induced changes in DMS reactivity *in vivo*. *a*, G ladders for both strands of a region of genomic DNA between −2,380 and −2,540 bp (upstream) of the transcriptional start site of the *TAT* gene from DMS-treated FTO-2B or XC cells (+, induced with dexamethasone; −, uninduced); *N*, *in vitro* methylated protein-free DNA. Protection (△) and enhancement (▲) seen when comparing dexamethasone-induced with uninduced FTO-2B cells; numbers, distance from the cap-site. Only those reactivity changes that appeared highly significant in the quantitative analysis (*c*) are indicated. *, Band originating from hybridization to *lacZ* sequences on a contaminating fragment; hatched bars, positions of the two major DNaseI footprints obtained *in vitro* with partially purified glucocorticoid receptor[17]. *b*, As in *a*, but only FTO-2B cells were used and incubation with the glucocorticoid antagonist RU486 (RU) was carried out in parallel with dexamethasone inductions. *c*, Quantitative analysis of hormone-dependent changes in G methylation; summary of results. Nucleotide numbers (from −2,560 to −2,390) upstream of the *TAT* cap-site are given on the abscissa. Hatched bars, sites of receptor binding identified by DNaseI footprinting *in vitro*[17]; black bar, conserved hexanucleotide TGTTCT. For each G residue the logarithm between the median of induced divided by the median of uninduced samples (log +dex/−dex) is plotted. △ and ▽ symbolize this value for each guanine on the upper and lower strand, respectively. Positive values indicate dexamethasone-dependent enhancements, negative values represent inducible protections. The significance of the deviations obtained was checked by comparing induced and uninduced values using a Kruskal-Wallis test[24]. Those values that fell below a threshold of 5% probability of error (1st order error) are indicated with vertical lines.

Methods. FTO-2B and XC cells were grown to about 80% confluency in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf-serum and 10 mM HEPES *p*H 7.4, withdrawn in serum-free DMEM for at least 16 h and incubated for 4 h with serum-free DMEM with or without 10^{-6} M dexamethasone or 10^{-6} M RU486. Induction times were 20 min–12 h, no significant time-dependent changes in the DMS reactivity pattern were observed (data not shown). DMS treatment of whole cells and *in vitro* methylation of protein-free DNA were as published[15,19]. After isolation of cellular DNA, 25 μg of each sample was restricted with *Hin*fI (Boehringer), reacted with piperidine and separated on a sequencing gel[12]. Details of DNA blotting onto Gene Screen membrane (NEN) UV crosslinking, hybridization and washing of filters were as in ref. 12. Probe synthesis from a single-stranded M13 template was performed as described[12] with modifications: 100 mM NaCls was included in primer annealing and elongation reactions, 250 μCi of [α^{32}P]dATP (NEN), 5,000 Ci mmole^{-1} was used per probe synthesis. The newly-synthesized probe DNA was separated from the template strand in a 6% denaturing gel and recovered from the polyacrylamide by isotachophoresis[22]. Exposure times were 7–14 days with 'lightning plus' intensifying screens (DuPont). Quantitative analysis was done as Giniger *et al.*[23]; four independent DMS treatments of FTO-2B dells were performed under either dexamethasone-induced or uninduced conditions. The resulting DNAs were analysed on at least two independent blots. Genomic sequencing ladders were scanned with an Elscript 400 densitometer (Fisher Scientific) at 16 μm resolution. Peak integrals were calculated and median values determined for each guanine band.

Fig. 3 Changes in DMS reactivity following glucocorticoid receptor binding *in vitro*. Left panel, results obtained for the upper strand and right panel, for the lower strand with (+) and without (−) triamcinolone-saturated glucocorticoid receptor. Protected guanines are indicated by open squares, enhanced methylation at a guanine as a filled square. Guanines that are influenced in their reactivity when comparing the '−' and '+' lanes are labelled with their numbers upstream of the *TAT* cap-site. Hatched bars, regions protected against DNaseI-digestion by bound glucocorticoid receptor *in vitro*[17].

Methods. Purification of the glucocorticoid receptor[25] included as a final step DEAE-Sepharose 6 CL column chromatography in which the receptor eluted as a single peak at 150 mM NaCl. A *Xba*I-*Tth*111I fragment containing both functional GREs was isolated from a subclone and end-labelled using T4 kinase polynucleotide and [γ^{32}P]ATP either at the *Xba*I (upper strand) or at the *Tth*111I (lower strand) sites. Labelled fragment (0.5–1.0 fmol) was incubated with (+) or without (−) 90 fmoles of triamcinolone-saturated receptor in a volume of 25 μl containing 100 mM NaCl, 0.6 mM EDTA, 7% glycerol (w/v), 22 mM Tris-Cl, 1 mM $MgCl_2$, 0.5 μg bovine serum albumin 6 mM β-mercaptoethanol, 0.1 mM dithiothreitol, 1 μM triamcinolone, 15% polyethyleneglycol 6000 (Serva), *p*H 8.2. After 45 min at room temperature, 2 μl of 10% DMS (Fluka) was added and reaction stopped after 2 min incubation at room temperature by adding 75 μl buffer containing 400 mM sodium acetate, 140 mM β-mercaptoethanol and 270 μg ml^{-1} transfer RNA. DNA was extracted with phenol/chloroform (1:1), precipitated and subjected to piperidine cleavage according to Maxam and Gilbert[19], and finally dissolved in formamide containing loading buffer and analysed on a 6% sequencing gel.



200 bp upstream of the transcriptional start site several protections and enhancements can be identified that are strictly correlated with the active state of the gene in FTO-2B cells but are not seen in XC cells. However, none of them are influenced by hormone treatment (P.B., unpublished data).

Comparison of the methylation patterns of the two analysed GREs *in vivo* reveals qualitative and quantitative differences. Whereas the protection of the central guanine (−2,499) in element II is clearly evident, it is less prominent in element III (−2,436). A surprising finding is a very strong protection of a TGT (−2,432) upstream of the conserved hexanucleotide TGTTCT of GRE III. Why the two binding sites behave so differently is not known. Recent results suggest that this difference in DMS reactivities might reflect differences in the

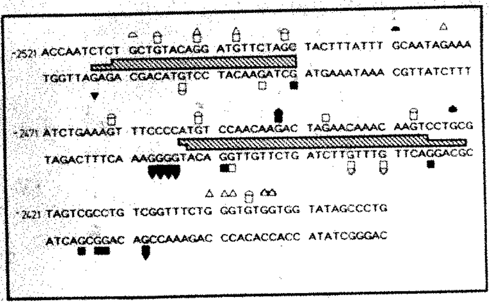

-2521 ACCAATCTCT GCTGTACAGG ATGTTCTAGC TACTTTATTT GCAATAGAAA
TGGTTAGAGA CGACATGTCC TACAAGATCG ATGAAATAAA CGTTATCTTT
-2471 ATCTGAAAGT TTCCCCATGT CCAACAAGAC TAGAACAAAC AAGTCCTGCG
TAGACTTTCA AAGGGGTACA GGTTGTTCTG ATCTTGTTTG TTCAGGACGC
-2421 TAGTCGCCTG TCGGTTTCTG GGTGTGGTGG TATAGCCCTG
ATCAGCGGAC AGCCAAAGAC CCACACCACC ATATCGGGAC

Fig. 4 Summary of results obtained by DMS reactivity experiments *in vivo* and *in vitro*. Changes in DMS reactivity at guanines in the DNA sequence between nucleotides −2,521 and −2,382, upstream of the *TAT* cap-site, are shown. Hatched bars, sites of glucocorticoid receptor interaction as identified by *in vitro* footprinting experiments[17]. Filled symbols, G residues showing enhanced methylation; open symbols, guanines showing protections from methylation. Squares, *in vitro* data; triangles, dexamethasone-dependent changes *in vivo*; half-circles, changes in reactivity found *in vivo* comparing the pattern obtained from uninduced cells with that from protein-free DNA. For the *in vivo* data only those guanines are indicated that by quantitative analysis have been shown to be significantly influenced in their reactivity (Fig. 2*c*).

inducing capacity of the two GREs (M. Jantzen and U.S., unpublished data).

In addition to effects in the receptor binding site, a number of changes in reactivity are observed in neighbouring sequences after dexamethasone treatment. Most prominent among these are the enhancements of guanines at −2,515, −2,459 to −2,456, and at −2,424 and protections of guanines at −2,475, −2,402 to −2,400, −2,396, and −2,395, *in vivo* and to some extent *in vitro*. Whether these changes are due to alterations in DNA structure as a consequence of glucocorticoid receptor binding or interaction with other factors not yet defined remains unclear.

These changes in DMS reactivity are clearly correlated with hormone induction and other changes are evident when the pattern of DMS reactivity of uninduced FTO-2B cell chromatin is compared with that of protein-free FTO-2B DNA (quantitative analysis not shown). Analysis of the reactivity of single guanines in naked genomic DNA at various sites and in plasmid DNA revealed that guanines flanked by thymidines or followed by a thymidine are often unusually sensitive to methylation. This enhanced reactivity is not observed *in vivo*, which might be explained by the difference in reaction conditions inside living cells compared with protein-free DNA *in vitro*. Because of these differences in methylation conditions naked genomic DNA cannot serve as a suitable standard for comparisons of DMS reactivity obtained *in vivo*. The methylation pattern in uninduced FTO-cells is identical to that obtained from induced or uninduced XC-cells. In contrast to observations at the cap-site, there is no evidence for an interaction of protein factors with the receptor binding site in the absence of hormone that discriminates *TAT* in FTO-cells from the gene in nonexpressing fibroblast cells.

We have applied genomic sequencing methodology to obtain information on the *in vivo* interaction of the glucocorticoid receptor with its recognition sequence in the *TAT* gene. Results support the concept that glucocorticoid increases the affinity of the receptor for its target sequence. These experiments do not exclude the possibility that the unliganded receptor can interact with a GRE, albeit with lower affinity and selectivity.

We thank Harry Nick for introducing us to the genomic sequencing methodology; Professor E. Weber for help with quantitative analysis, Alfred Nordheim, Richard Miksicek, Francis Stewart and Michael Boshart for critical reading of the manuscript and M. Cole for secretarial assistance. This work was supported by the Deutsche Forschungsgemeinschaft (Schu 51/4-2).

Received 23 June; accepted 28 October 1986.

1. Payvar, F. P. *et al. Cell* **35**, 381-392 (1983).
2. Scheidereit, C., Geisse, S., Westphal, H. M. & Beato, M. *Nature* **304**, 749-752 (1983).
3. Van der Ahe, D. *et al. Nature* **313**, 706-709 (1985).
4. Yamamoto, K. R. *A. Rev. Genet.* **19**, 209-252 (1985).
5. Hynes, N. *et. al. Proc. natn. Acad. Sci. U.S.A.* **80**, 3637-3641 (1983).
6. Renkawitz, R., Schütz, G., von der Ahe, D. & Beato, M. *Cell* **37**, 503-510 (1984).
7. Karin, M. *et al. Nature* **308**, 513-519 (1984).
8. Miksicek, R. *et al. Cell* **46**, 283-290 (1986).
9. King, W. J. & Greene, G. L. *Nature* **307**, 745-747 (1984).
10. Welshons, W. K., Liebermann, M. E. & Gorski, J. *Nature* **307**, 747-749 (1984).
11. Perrot-Applanat, M., Logeat, F., Groyer-Picard, M. T. & Milgrom, E. *Endocrinology* **116**, 1473-1484 (1985).
12. Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991-1995 (1984).
13. Church, G. M., Ephrussi, A., Gilbert, W. & Tonegawa, S. *Nature* **313**, 798-801 (1985).
14. Nick, H. & Gilbert, W. *Nature* **313**, 795-797 (1985).
15. Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. *Science* **227**, 134-140 (1985).
16. Becker, P. B., Renkawitz, R. & Schütz, G. *EMBO J.* **3**, 2015-2020 (1984).
17. Jantzen, M. *et al.* (in preparation).
18. Zaret, K. S. & Yamamoto, K. R. *Cell* **38**, 29-38 (1984).
19. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499 (1978).
20. Ogata, R. T. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5851-5854 (1978).
21. Scheidereit, C. & Beato, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3029-3033 (1984).
22. Öfverstedt, L. G. *et al. Biochim. biophys. Acta* **782**, 120-126 (1984).
23. Gininger, E., Varnum, S. M. & Ptashne, M. *Cell* **40**, 767-774 (1985).
24. Kruskal, W. H. *Ann. math. Statist.* **23**, 525-540 (1952).
25. Singh, V. B. & Moudgil, V. K. *J. biol. Chem.* **260**, 3684-3690 (1985).

Steroid-free glucocorticoid receptor binds specifically to mouse mammary tumour virus DNA

Toivo Willmann & Miguel Beato*

Institut für Molekularbiologie und Tumorforschung,
Emil-Mannkopff-Strasse 1, 3550 Marburg, FRG

Steroid hormones are thought to modulate gene expression through their interaction with receptor proteins. The intracellular localization of unoccupied receptor proteins has been a subject of controversy: free glucocorticoid receptor appears to reside in the cytoplasm and moves to the cell nucleus only after binding the steroid[1,2]. The purified hormone-bound glucocorticoid receptor has been shown to bind selectively to hormone regulatory elements (HRE) in the vicinity of hormonally-inducible promoters and, in particular, in the long terminal repeat (LTR) region of mouse mammary tumour virus (MMTV)[3,4]. We have tackled the question of whether the hormone itself is required for the interaction of the receptor protein with the HRE. Using monoclonal antibodies to the receptor we find that upon heat-activation the steroid-free glucocorticoid receptor present in rat liver cytosol binds specifically *in vitro* to the HRE of MMTV. No qualitative differences in the DNaseI-footprints were detected when hormone-free receptor was compared to the hormone-receptor complex or even receptor complexed with the hormone antagonist RU486. We conclude that the steroid ligand is not an absolute requirement for generating the conformation of the glucocorticoid receptor that allows its interaction with the HRE *in vitro*. An alternative function of the hormone *in vivo* could be to modulate nuclear partitioning of the receptor.

To determine whether the steroid ligand is required for specific binding of the receptor to DNA it is important to use receptor that has been subjected to as few manipulations as possible, because the steps required for receptor purification could induce artificial conformational changes of the receptor protein, and thus alter its behaviour in DNA-binding studies. Therefore, we decided to use the receptor in crude hepatic cytosol from adrenalectomized rats. To show that the glucocorticoid receptor in this cytosol is free of ligand we used a competition assay. Cytosol was heated at 100 °C for 4 min to denature the steroid-binding proteins and then extracted with *n*-butanol. The amount of ligand with affinity for the glucocorticoid receptor in the

---

[*] To whom correspondence should be addressed.

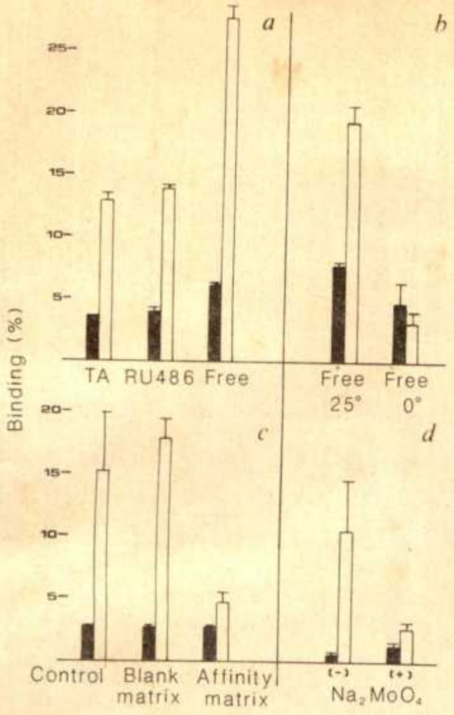


Fig. 1 Immunological detection of DNA-receptor complexes. *a*, DNA-cellulose filtered hepatic cytosol from adrenalectomized rats, either charged with 50 nM [^{3}H]triamcinolone acetonide (TA) or 100 nM RU468 (RU468) or free of hormones (FREE), was activated at 25 °C and aliquots incubated with radioactive DNA fragments derived from plasmid p13-13[6], which contains parts of the *env* gene and the LTR region of MMTV. Monoclonal antireceptor antibodies were added and after incubation precipitated with a second antibody linked to Pansorbin. Bound radioactivity was counted. The average and variation in duplicate measurements are shown. Black bars, binding of a 2.7 kb control fragment (vector sequences and part of the *env* gene), not selectively bound by purified glucocorticoid receptor[4]. Open bars, binding of a 2.3 kb fragment containing the LTR region of MMTV together with mouse genomic sequences and the herpes simplex virus (HSV) *tk* gene. *b*, Left, as in *a* with hormone-free cytosol (25 °C); right, the same but performed at 0 °C during all the incubations to prevent thermal activation (0 °C). *c*, Experiment with uncharged cytosol preincubated at 4 °C for 16 h in presence of 15 vol% of pure Sepharose 4B (Blank matrix), Sepharose 4B containing 0.5 mg ml^{-1} of immobilized dexamathasone (Affinity matrix) or in absence of Sepharose (Control). Shorter DNA fragments were used: black bars, 450 base pairs (bp), *env* and vector sequences; open bars, 600 bp, MMTV-LTR and mouse genomic sequences. *d*, Hormone-free cytosol activated and incubated in the absence (−) or the presence (+) of 20 mM $NaMoO_4$ and with DNA fragments as in *c*.

Methods. Cytosol was passed through DNA-cellulose and receptor content determined by the charcoal technique[5]. Aliquots containing 10 ng of receptor with or without steroid ligand were heat activated at 25 °C for 30 min and incubated for additional 30 min with 10 ng of the end-labelled restriction fragment and 30 ng (with the large restriction fragments, Fig. 1*a*, *b*) or 1 μg (with the short fragments, Fig. 1*c*, *d*) of calf thymus DNA (sheared to an average size of 0.5 kb kilobase (kb) by treatment with 0.3 M NaOH at 100 °C for 30 min) in a final volume of 50 μl containing: 10 mM Tris HCl *p*H 7.5, 0.1 mMm Na_2 EDTA & 110 mM NaCl. After the addition of 50 ng of Protein-A-Sepharose purified monoclonal antireceptor antibody IGR49/4 the incubation was continued at 25 °C for 15 min. The assays were rotated at 4 °C with 1 μl of solid Pansorbin containing 50 μg of purified rabbit anti mouse IgG in a final volume of 200 μl containing: 12 mM HEPES, *p*H 6.8, 150 mM NaCl and at 0.5–1 μg calf thymus DNA. After centrifugation the pellets were washed with Tris-buffer containing 1 mg ml^{-1} bovine serum albumin and 110 mM NaCl.

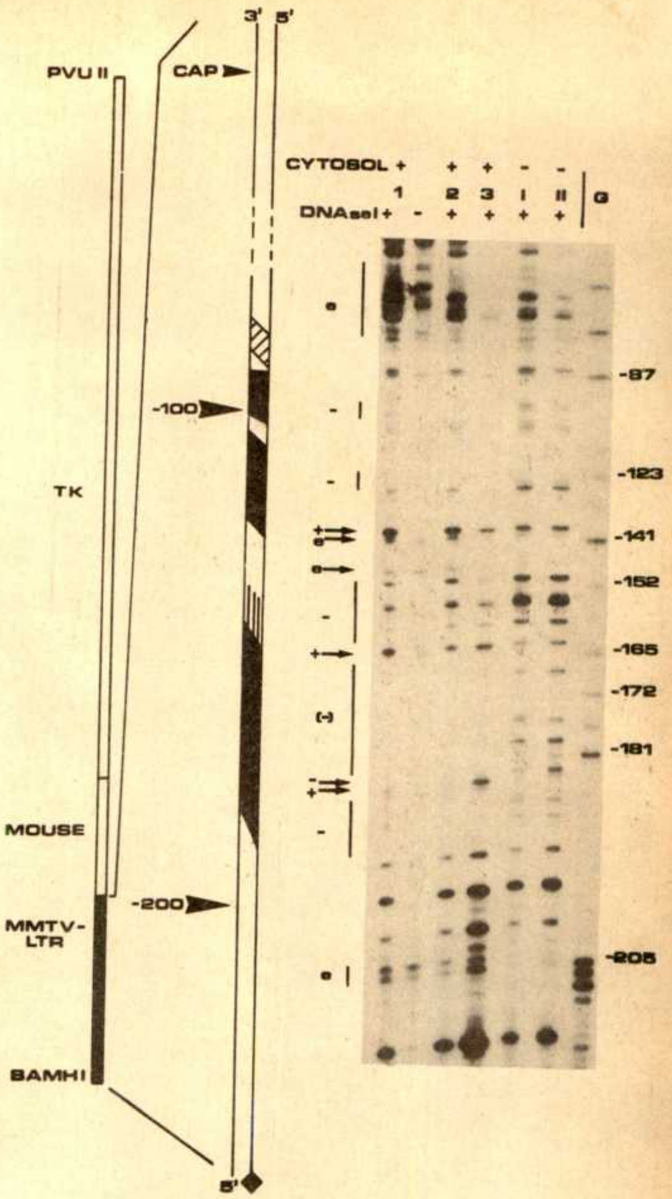


Fig. 2 DNaseI footprints with crude cytosol receptors. DNaseI footprints generated by the receptor complexed with corticosterone (lane 1+), with RU468 (lane 3) or free of hormone (lane 2). Controls: Lane 1−, like lane 1+ but without added DNaseI; lanes I, and II, DNaseI digestions in absence of cytosol. Lane G: guanosine-specific sequencing reaction[19]. The fragment used and position of the label (◆) are shown at the left. −, regions protected by the receptor; +, hypersensitive ones; e, cuts by endogenous nucleases (lane 1−). Numbers refer to the distance from the initiation of transcription (CAP).

Methods. The 2.3 kb fragment of p13-13 was 3′-labelled using dGTP, [α-^{32}P]dATP and the Klenow fragment of DNA-polymerase I. It was used in an assay like that in Fig. 1*a*. The precipitate was resuspended in 80 μl of washing buffer and calf thymus (1 μg) and $MgCl_2$ (final concentration 2.3 mM) in 20 μl were added. After the addition of 0.1 μg DNaseI (Worthington) and incubation at 20 °C for 60 sec the digestion was stopped with 3 μl 250 mM Na_2 EDTA. The DNA was solubilized with 0.6% SDS and 0.3 M NaCl and purified by cooling to 0 °C, centrifugation, fivefold extraction with chloroform/isoamylalcohol and precipitation. After digestion with Pronase P the samples were analysed in a 6.5% polyacrylamide-urea sequencing gel[19]. For lanes I and II 30 ng of labelled DNA were digested in an assay volume of 300 μl with 0.4 μg (I) or 0.9 μg (II) DNaseI.

Table 1 Glucocorticoid receptor ligand in cytosol from adrenalectomized rats

Competitor added		Binding of [^{3}H] triamcinolone acetonide	
Corticosterone	*n*-Butanol extract	c.p.m.	%
—	—	3,215 (154)	100
1 ng	—	1,954 (2)	61
2 ng	—	1,418 (74)	44
5 ng	—	865 (12)	27
25 ng	—	412 (8)	13
—	2 ml	3,468 (247)	108
—	2 ml*	3,140	98

Aliquots of rat liver cytosol from adrenalectomized rats were incubated with [^{3}H]triamcinolone acetonide in the presence of corticosterone or *n*-butanol extract from 2 ml of cytosol. At equilibrium the protein bound radioactivity was determined by the charcoal technique. Values are average and variation of duplicate determinations[5]. To obtain cytosol, male Wistar rats (body weight 180–200 g) were adrenalectomized, maintained under standard conditions with physiological saline *ad libitum* and killed after four days. The livers were perfused through the portal vein with 20 ml cold homogenization buffer (20 mM sodium phosphate, *p*H 7.0, 1 mM Na$_2$EDTA, 50 mM NaCl, 2 mM β-mercaptoethanol, 2 mM dithioerythritol and 10% glycerol), removed, minced and homogenized in three volumes of the same buffer. After centrifugation at 109,000 *g* and 4 °C for 90 min, the supernatant was passed through 0.14 volumes of calf-thymus-DNA-cellulose (0.5 mg DNA per g cellulose). The flow-through was collected and stored in liquid nitrogen. For *butanol extracts*, 2 ml aliquots of cytosol were heated at 100 °C for 4 min, extracted with one volume *n*-butanol by vortexing for 30 s and concentrated by evaporation. A control extraction with charcoal-treated cytosol supplemented with [^{3}H]cortisol (200 ng, 0 °C for 60 min) showed that this procedure extracts 86% of the hormone. For the competition assay, cytosol was treated with 10% dextran-coated charcoal[5] at 0 °C for 5 min, and aliquots (100 μl) were preincubated for 5 min in the presence of either corticosterone or *n*-butanol extract, before addition of 0.9 ng of [^{3}H]triamcinolone acetonide (Amersham Büchler, specific activity 1.1 TBq mmol^{-1}) and further incubation for 60 min at 0–4 °C. The amount of protein-bound radioactivity was determined by the charcoal procedure[5].

* In this experiment the *n*-butanol extract was added only 5 min prior to the charcoal treatment to exclude an effect of the organic extract *per se* on the efficiency of the test.

organic extract was determined by measuring its ability to compete with ^{3}H triamcinolone acetonide for binding to liver cytosol treated with charcoal[5]. In the hepatic cytosol we used there is less than 0.5 ng ml^{-1} of ligand, so that given the association constant of the receptor for corticosterone, at least 98% of the receptor molecules are free of steroid (Table 1). The unoccupied receptor molecules are able to bind steroid, as shown by their retention in a steroid affinity matrix (Fig. 1*c*, below).

For DNA-binding experiments we used two end-labelled restriction fragments of approximately equal size derived from a plasmid carrying the LTR region of MMTV (Fig. 1, ref. 6). The individual fragments were incubated with hormone-free cytosol or with cytosol that was precharged with either the synthetic glucocorticoid triamcinoline acetonide or the potent antagonist RU486[7]. In every case the cytosol was incubated at 25 °C for 30 min before the addition of DNA. The complexes of receptor and DNA were immunoprecipitated with mouse monoclonal antibodies against the receptor, IGR49/4[8], and rabbit antimouse immunoglobulin linked to Pansorbin (Calbiochem). The radioactivity in the precipitates was measured and is expressed as percentage of the total radioactive DNA. In all experiments the restriction fragment containing LTR sequences was immunoprecipitated three to four times more efficiently than the slightly larger vector fragment (Fig. 1*a*). The hormone-free receptor binds specifically to the LTR-fragment, and exhibits a total DNA-binding activity higher than that of the cytosol incubated with either receptor ligand. This result is not influenced by pre-treatment of the cytosol with dextran-coated charcoal (data not shown), nor by using shorter restriction fragments that increase the specificity of DNA binding (Fig. 1*c*, *d*). No significant differences were found between the DNA binding properties of cytosol charged with agonist or antagonist ligands (Fig. 1*a*).

Specific binding to the LTR fragment was dependent upon previous heat activation of the hormone-free cytosol and could be prevented by addition of 20 mM sodium molybdate (Fig. 1*b*, *d*). Binding of the antibody *per se* does not activate the receptor by artificial conformational change (Fig. 1*b*). That the receptor molecules responsible for specific DNA binding in hormone-free cytosol are able to bind steroid was shown by passing the cytosol through an affinity matrix of immobilized dexamethasone[9]. This treatment removes from the cytosol 85% of the steroid-binding activity, eliminates any immunoreactive material in Western blots (data not shown), and dramatically decreases specific DNA binding (Fig. 1*c*). Thus, a steroid-free form of the glucocorticoid receptor is able to bind active hormone and interacts selectively with the LTR of MMTV. These results were confirmed by a nitrocellulose filter binding assay which did not depend on the use of antibodies (data not shown).

To analyse in more detail the interaction of the different receptor forms with the LTR region we performed DNase I footprinting experiments with the immunoisolated protein–DNA complexes. Though the quality of the footprints is not comparable to those obtained with purified receptor, the limits of the region protected against DNase I are the same as those previously reported[4], and do not differ in the footprints obtained with hormone-free cytosol or with liganded cytosols (Fig. 2). Small quantitative differences between the protection in the presence of agonist and antagonist ligands can not be excluded by these results.

We conclude that the hormone-free glucocorticoid receptor in crude rat liver cytosol binds *in vitro* to the same regulatory elements of the MMTV-LTR region that have been shown to mediate hormonal regulation of transcription *in vivo*[3,4,6]. Our results apparently contradict previous reports, based on binding of radioactive steroid-receptor complexes to an excess of calf thymus DNA-cellulose[10,11], that binding of the hormone to the receptor is a prerequisite for the complex to be 'activated' and bind to DNA. We have repeated these experiments and find that the steroid-free receptor can be heat activated and bind to DNA-cellulose, though the extent of binding is only 20–40% of that observed with the hormone-receptor complexes (data not shown). One possible reason for the quantitative discrepancy with the results presented here could be that an excess of calf thymus DNA-cellulose acts as an ionic exchanger rather than a specific ligand for the DNA-binding domain of the receptor. This could also explain the difference between our results and a previous report claiming that, in a DNA-cellulose competition assay, the receptor complexed to the antagonist RU486 binds less efficiently to MMTV-LTR than receptor complexed to agonist[12].

Our results confirm the reequirement for an *in vitro* heat activation step in order for the receptor, whether free or steroid-bound, to bind to the DNA regulatory sequences. At physiological body temperatures receptor activation *in vivo* should be spontaneous. In the absence of hormone the receptor seems to be localized in the cytoplasm, so that some mechanism preventing receptor activity must exist *in vivo*.

It is conceivable that in the absence of hormone the receptor is complexed to other cellular factors or structures that prevent its interaction with the physiological nuclear targets. Candidates for this role are RNA and the 90 kilodalton heat shock protein, which have been found associated with the non-activated form of several steroid receptors[13,14]. The function of the hormone could be to dissociate the receptor from the 'anchoring' structure, and thus to allow its translocation to the cell nucleus. As a novel variation of this concept, the actual function of the hormone may be to induce a conformational change of the receptor that exposes a nuclear translocation signal. Amino acid sequences responsible for nuclear translocation have been identified in several nuclear proteins, and appear to be distinct from the actual DNA-binding site[15,16]. In the absence of the hormone this hypothetical nuclear translocation signal would be masked and the receptor protein would be excluded from the cell nucleus. The possible role of receptor phosphorylation in this remains an intriguing question.

Common to all of these models is the concept that the DNA binding domain of the receptor is functional in the absence of hormone. In the case of oestrogens there is evidence supporting a weak association of the steroid-free receptor with nuclear structures[17,18]. Whether the hormone ligand is required for the activation of gene expression that follows DNA binding could be analysed in appropriate cell-free experiments.

We thank Hannes M. Westphal for the monoclonal receptor antibodies and Pierre Formstecher for the dexamethasone affinity matrix. This work was supported by grants from the Deutsche Forschungsgemeinschaft and from the Fond der Chemischen Industrie.

Received 27 May; accepted 23 October 1986.

1. Fuxe, K. *et al. Endocrinology* **117,** 1803–1812 (1985).
2. Antakly, T. & Eisen, H. J. *Endocrinology* **115,** 1984–1989 (1984).
3. Payvar, F. *et al. Cell* **35,** 381–392 (1983).
4. Scheidereit, C., Geisse, S., Westphal, H. M. & Beato, M. *Nature* **304,** 749–752 (1983).
5. Beato, M. & Feigelson, P. *J. biol. Chem.* **247,** 7890–7896 (1972).
6. Hynes, N. H. *et al. Proc. natn. Sci. U.S.A.* **80,** 3637–3641 (1983).
7. Herrmann, W. *et al. C.r. Acad. Sci. Paris* **294,** (III), 933–938 (1982).
8. Westphal, H. M., Moldenhauer, G. & Beato, M. *EMBO J.* **1,** 1467–1471 (1982).
9. Idziorek, T. *et al. Eur. J. Biochem.* **153,** 65–74 (1985).
10. Baxter, J. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **69,** 1892–1897 (1972).
11. Beato, M., Kalimi, M., Konstam, M. & Feigelson, P. *Biochemistry* **12,** 3372–3379 (1973).
12. Bourgeois, S., Pfahl, M. & Baulieu, E. E. *EMBO J.* **3,** 751–755 (1984).
13. Joab, I. *et al. Nature* **308,** 850–853 (1984).
14. Tymoczko, J. L. & Phillips, M. *Endocrinology* **112,** 142–149 (1983).
15. Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. *Cell* **39,** 499–509 (1984).
16. Richardson, W. D., Roberts, B. L. & Smith, A. E. *Cell* **44,** 77–85 (1986).
17. King, W. J. & Greene, G. L. *Nature* **307,** 745–747 (1984).
18. Welshons, W. V., Liebermann, M. E. & Gorski, J. *Nature* **307,** 747–749 (1984).
19. Maxam, A. M. & Gilbert, W. *Meth. Enzymol.* **5,** 499–560 (1980).

Molecular cloning and polymorphism of the human immune deficiency virus type 2

François Clavel, Mireille Guyader, Denise Guétard, Mireille Sallé*, Luc Montagnier & Marc Alizon†

Unité d'Oncologie virale, and CNRS UA1157, Institut Pasteur,
25 rue du Dr Roux, 75724 Paris CEDEX 15, France
* Diagnostics Pasteur, 3 Bd. R. Poincaré, 92350 Marnes, France

We recently reported the isolation of a novel retrovirus, the human immune deficiency virus type 2 (HIV-2, previously named LAV-2), from patients with acquired immune deficiency syndrome (AIDS) originating from West Africa[1,2]. This virus is related to HIV-1, the causative agent of the AIDS epidemic now spreading in Central and East Africa, as well as the USA and Europe (see ref. 3 for review) both by its morphology and by its tropism and *in vitro* cytopathic effect on CD4 (T4) positive cell lines and lymphocytes. But preliminary hybridization experiments indicated that there are substantiated differences between the sequences of the two genomes[2]. Furthermore, the proteins of HIV-1 and HIV-2 have different sizes and their serological cross-reactivity is restricted to the major core protein, as the envelope glycoproteins of HIV-2 are not immunoprecipitated by HIV-1-positive sera[1,2]. We now report the molecular cloning of the complete 9.5-kilobase (kb) genome of HIV-2, the observation of restriction site polymorphism between different isolates, and a preliminary analysis of the relationship of HIV-2 with other human and simian retroviruses.

A series of filter hybridizations of the HIV-2 RNA genome with probes derived from the complete cloned HIV-1 genome, as well as from regions expected to be well-conserved between HIV-1 and 2 (principally from the *gag* and *pol* genes) yielded extremely weak signals, even in conditions of very low stringency hybridization and washing[2]. This made it very difficult to assess by Southern blot analysis the amount of HIV-2 viral and proviral DNA in infected cells. We therefore cloned a complementary DNA (cDNA) to the HIV-2 genomic RNA, to produce a specific hybridization probe. We used the strategy previously developed for the cloning of HIV-1 (ref. 4). An oligo(dT)-primed cDNA first strand was made in the detergent-activated endogenous reaction (using HIV-2 reverse transcriptase) with virions purified from supernatants of CEM cells (a lymphoblastoid $CD4^+$ cell line[5]) infected with the isolate ROD[2] and continuously producing high amounts of HIV-2. A collection of 10^4 M13 recombinant phages was obtained and screened *in situ* with a labelled HIV-2 complementary DNA first strand. As previously done for HIV-1, we looked directly for HIV-2 specific clones, expected to represent an abundant group in the library, but detected only clones corresponding to cellular repetitive DNA. We can now explain this result by the low representation of HIV-2 specific clones in the library (less than 0.5%, compared to the 5–10% obtained in the cloning of HIV-1). Fortunately the high amounts of recombinant DNA produced by M13, its single-stranded nature, and the ability to perform extremely low stringency hybridization (at T_m −42 °C) on replica filters of M13 plaques allowed the successful use of an HIV-1 probe spanning 1.5 kb of the 3′ end of the LAV_{BRU} isolate (Fig. 1*a*) to screen the library. About 50 positive plaques were detected, purified and characterized by end sequencing and cross-hybridizing of the inserts (Fig. 1). The different clones were found to be complementary to the 3′ end of a polyadenylated RNA, having the AATAAA signal about 20 nucleotides upstream of the poly(A) tail, as found in the long terminal repeat (LTR) of HIV-1 (ref. 6). The part of the HIV-2 LTR that we sequenced is only very distantly related to the homologous domain in HIV-1 (Fig. 1*b*). Only ~50% of the nucleotides can be aligned, requiring the introduction of ~100 insertions and deletions. For comparison, the homology of the corresponding domains in HIV-I isolates from USA and Africa is greater than 95%, with no insertions or deletions (M.A., unpublished data).

The largest insert of this group of M13 clones (E2, 2 kb) was used as a probe to demonstrate HIV-2 specificity in a series of filter hybridization experiments. First, this probe could detect the genomic RNA of HIV-2 but not HIV-1 in stringent conditions (Fig. 2*C*,*D*). Second, positive signals were detected in Southern blots of DNA from cells infected with the ROD isolate as well as other isolates of HIV-2 (Fig. 2*A* and Fig. 4*A*). No signal was detected with DNA from uninfected (not shown), or HIV-1 infected cells (Fig. 2*A*) confirming the exogenous nature of HIV-2. In undigested DNA from HIV-2 infected cells we detect principally a species of ~10 kb, probably corresponding to linear unintegrated viral DNA, and a species with an apparent size of 6 kb, probably the circular form of the viral DNA. Conversely, rehybridization of the same filter with an HIV-1 probe under stringent conditions showed hybridization to HIV-1 infected cells only (Fig. 2*B*).

To isolate the rest of the genome of HIV-2, a genomic library in lambda phage L47.1 (ref. 7) was constructed with a partial *Sau*3AI restriction digest of the DNA from the CEM cell line infected with $HIV\text{-}2_{ROD}$. About 2×10^6 recombinant plaques

---

† To whom correspondence should be addressed.

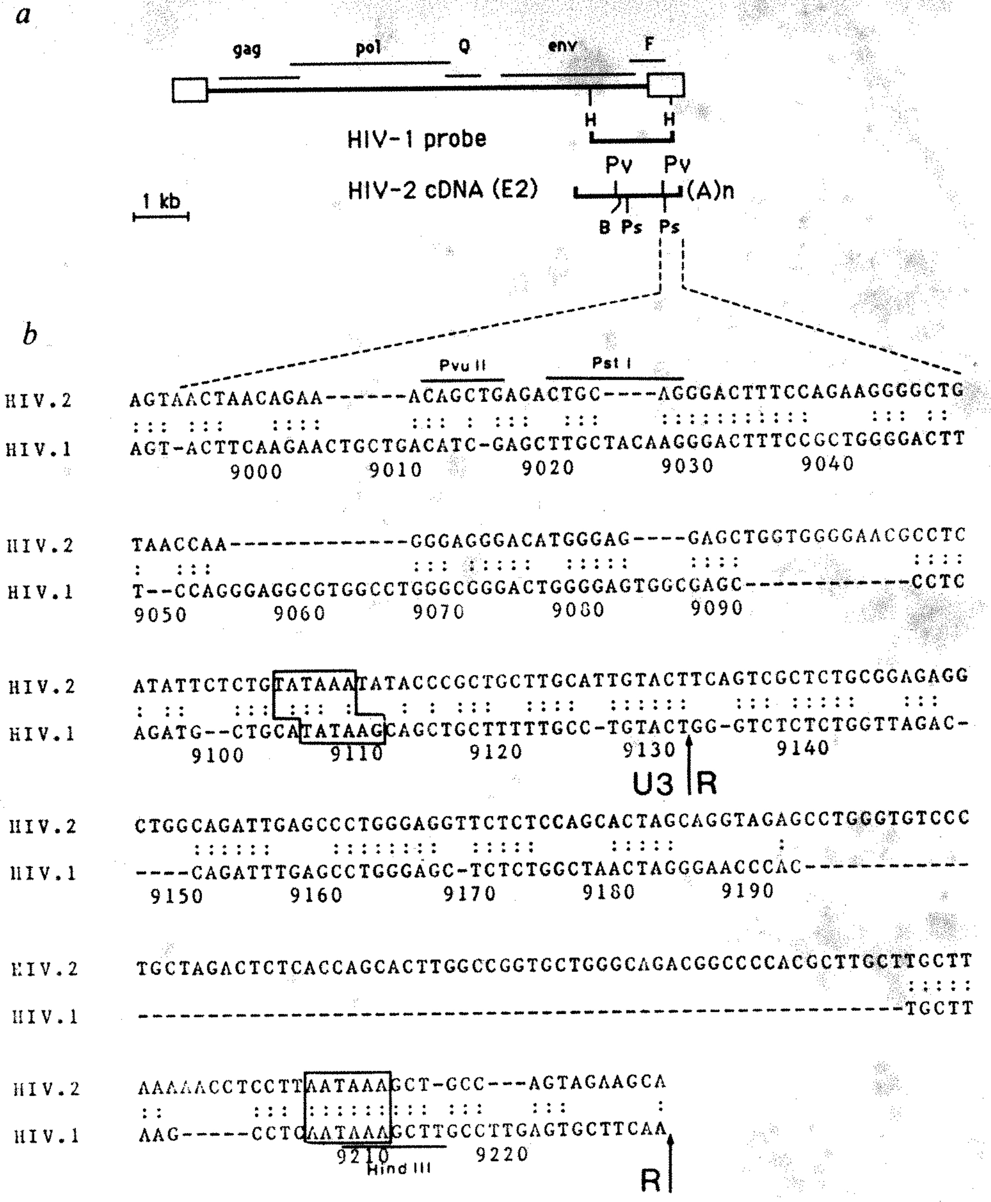

a
gag
pol
Q
env
F
H
H
HIV-1 probe
Pv
Pv
HIV-2 cDNA (E2)
(A)n
1 kb
B
Ps
Ps
b
Pvu II
Pst I
HIV.2 AGTAACTAACAGAA------ACAGCTGAGACTGC----AGGGACTTTCCAGAAGGGGCTG
HIV.1 AGT-ACTTCAAGAACTGCTGACATC-GAGCTTGCTACAAGGGACTTTCCGCTGGGGACTT
9000 9010 9020 9030 9040
HIV.2 TAACCAA-------------GGGAGGGACATGGGAG----GAGCTGGTGGGGAACGCCTC
HIV.1 T--CCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGC-----------CCTC
9050 9060 9070 9080 9090
HIV.2 ATATTCTCTGTATAAATATACCCGCTGCTTGCATTGTACTTCAGTCGCTCTGCGGAGAGG
HIV.1 AGATG--CTGCATATAAGCAGCTGCTTTTTGCC-TGTACTGG-GTCTCTCTGGTTAGAC-
9100 9110 9120 9130 9140
U3
R
HIV.2 CTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAGCAGGTAGAGCCTGGGTGTCCC
HIV.1 ----CAGATTTGAGCCTGGGAGC-TCTCTGGCTAACTAGGGAACCCAC-----------
9150 9160 9170 9180 9190
HIV.2 TGCTAGACTCTCACCAGCACTTGGCCGGTGCTGGGCAGACGGCCCCACGCTTGCTTGCTT
HIV.1 ---TGCTT
HIV.2 AAAAACCTCCTTAATAAAGCT-GCC---AGTAGAAGCA
HIV.1 AAG-----CCTCAATAAAGCTTGCCTTGAGTGCTTCAA
9210 9220
Hind III
R

Fig. 1 Cloned complementary DNA (cDNA) from the genomic RNA of HIV-2; *a*, genetic organization of HIV-1, position of the HIV-1 *Hin*dIII fragment used as a probe to screen the cDNA library, and restriction map of the HIV-2 cDNA clone, E2. B, *Bam*HI; H, *Hin*dIII; Ps, *Pst*I; Pv, *Pvu*II. *b*, partial nucleotide sequence of the LTR of HIV-2. The corresponding region of the HIV-1 LTR[6] was aligned using the Wilbur and Lipman algorithm (window, 10; K-tuple, 7; gap penalty, 3; ref. 14). The U3-R junction in HIV-1 is indicated; poly(A) addition signal and potential TATA promoter regions are boxed.

Methods. HIV-2 virions were purified from 5 litres of supernatant from a culture of the CEM cell line infected with the ROD isolate[2] and a cDNA first strand using oligo(dT) primer was synthesized in a detergent-activated endogenous reaction on pelleted virus resuspended in 1 ml, as described[4]. RNA-cDNA hybrids were purified by phenol-chloroform extraction, and ethanol precipitation. The second-strand cDNA was done by the DNA polymerase I/RNAase H method[15] and blunt-ended with T4 DNA polymerase using a commercial cDNA synthesis kit (Amersham). After attachment of *Eco*RI linkers (Pharmacia), *Eco*RI digestion, and ligation into *Eco*RI-digested dephosphorylated M13 *tg*130 vector (Amersham), a cDNA library was obtained by transformation of the *Escherichia coli* TG1 strain. Recombinant plaques (10^4) were screened *in situ* on replica filters with the 1.5-kb *Hin*dIII fragment from clone J19, corresponding to the 3′ part of the genome of the LAV$_{BRU}$ isolate of HIV-1 (ref. 4), ^{32}P-labelled to a specific activity of 10^9 c.p.m. μg^{-1}. The filters were prehybridized in 5 × SSC, 5 × Denhardt solution, 25% formamide, denatured salmon sperm DNA (100 $\mu g\ ml^{-1}$) at 37 °C, for 4 h and hybridized for 16 h in the same buffer ($T_m - 42$ °C) plus 4×10^7 c.p.m. of the labelled probe (10^6 c.p.m. per ml of hybridization buffer). The washing was done in 5 × SSC, 0.1% SDS at 25 °C for 2 h. 20 × SSC is 3 M NaCl, 0.3 M Na citrate. Positive plaques were purified and single-stranded M13 DNA prepared and end-sequenced[16].

were screened *in situ* with labelled insert from the E2 cDNA clone. Ten recombinant phages were detected and plaque-purified; three of them were characterized by restriction mapping and Southern blot hybridization with the E2 insert and probes from its 3′ end (LTR) or 5′ end (envelope), as well as with HIV-1 subgenomic probes (HIV-1 probes were used under non-stringent conditions). The λ ROD 4 clone is very likely to contain the complete genetic information of HIV-2, as it carries a 9.5-kb insert and is derived from a circular viral DNA. Two other clones, λ ROD 27 and λ ROD 35 are derived from integrated proviruses and carry a LTR and cellular flanking sequences and a portion of the viral coding sequences (Fig. 3*A*).

The relationship of HIV-2 to other human and simian retroviruses will be more precisely determined by the elucidation of the nucleotide sequence of their genomes, but some preliminary conclusions can already be drawn from hybridization experiments. The relative homology of the different regions of the HIV-1 and 2 genomes was determined by hybridization of fragments of the cloned HIV-1 genome with the labelled λROD 4 expected to contain the complete HIV-2 genome (Fig. 3*B*). Even at very low stringency ($T_m - 42$ °C), the hybridization of HIV-1 and 2 is restricted to a fraction of their genomes, principally the *gag* gene (dots 1 and 2), the reverse transcriptase domain in *pol* (dot 3), the end of *pol* and the Q (or *sor*) genes (dot 5) and the F gene (or 3′ ORF) and 3′ LTR (dot 11). The HIV-1 fragment used to detect the HIV-2 cDNA clones contains the dot 11 subclone, which indeed hybridizes well to HIV-2 under non-stringent conditions. Only the signal from dot 5 persists after stringent washing. The envelope gene, but also the region of the *tat* gene and a part of *pol* seem very very divergent. These data, and also the LTR sequeence (Fig. 1*B*), indicate that HIV-2 is not an envelope variant of HIV-1.

We previously observed that HIV-2 is more related to the simian immune deficiency virus (SIV), than it is to HIV-1 (ref. 2). SIV (also designated simian T-cell lymphotropic virus type 3, STLV-3) is a retrovirus first isolated from captive macaques with an AIDS-like disease in the USA[8]. All the SIV proteins, including the envelope, are immunoprecipitated by sera from HIV-2

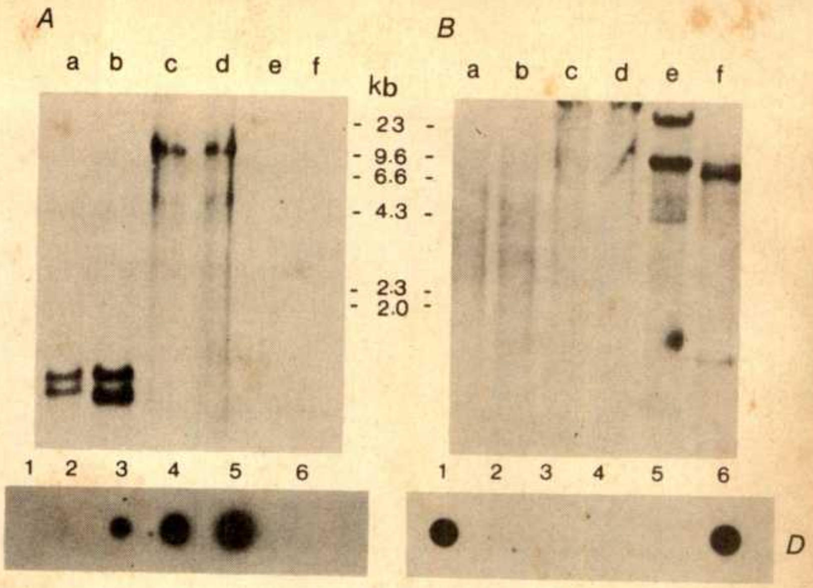


Fig. 2 HIV-2 specificity of the E2 cDNA clone. *A*, *B*, Southern blots of DNA extracted from CEM cells infected with isolates: HIV-2_{ROD} (a, c) HIV-2_{DUL} (b, d) and HIV-1_{BRU} (e, f); a, b, f, *Pst*I digested; c, d, e, undigested. *C*, *D*, dot blot hybridization of pelleted virions from CEM cells infected by the HIV-1_{BRU} (1), SIV isolate Mm 142-83 (3), HIV-2_{DUL} (4), HIV-2_{ROD} (5), HIV-1_{ELI} (6). Dot 2 is a pellet from an equivalent volume of supernatant from uninfected CEM. *A*, *C*, Hybridization with the HIV-2 cDNA (E2); *B*, *D*, hybridization to HIV-1 probe (*Sac*I 9-kb insert from HIV-1_{BRU}, ref. 4).

Methods. DNA was extracted from infected CEM cells continuously producing HIV-1 or 2, and 20 μg of *Pst*I-digested or undigested DNA, was electrophoresed on a 0.8% agarose gel, and transferred to nylon membrane. Virion dot blots were prepared in duplicate as described[2], by pelleting volumes of supernatant corresponding to the same amount of reverse transcriptase activity. Prehybridization was done in 50% formamide, 5 × SSC, 5 × Denhardt's solution, 100 μg ml^{-1} denatured salmon sperm DNA for 4 h at 42 °C. Hybridization was performed in the same buffer plus 10% Dextran sulphate, and 10^6 c.p.m. per ml of the labelled E2 insert (specific activity 10^9 c.p.m. μg^{-1}) for 16 h at 42 °C. Washing was in 0.1 × SSC, 0.1% SDS for 2 × 30 minutes. After exposition (16 h with intensifying screens), the Southern blot was dehybridized in 0.4 M NaOH, neutralized, and rehybridized in the same conditions to the HIV-1 probe labelled to 10^9 c.p.m. μg^{-1}.

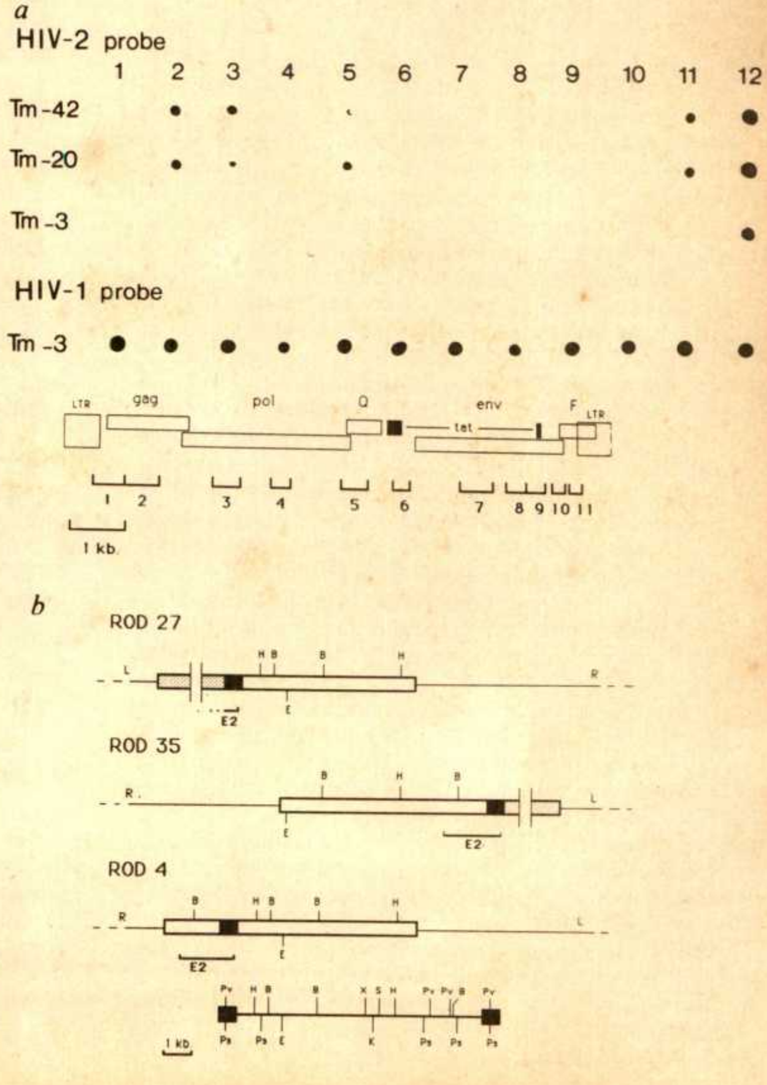


Fig. 3 Restriction map of the HIV-2_{ROD} genomes and similarity to HIV-1. *A*, Organization of three recombinant phage λ clones, ROD 4, ROD 27 and ROD 35; the open boxes represent viral sequence (the LTR are filled), whereas the dotted boxes represent cellular flanking sequences (not mapped); only some characteristic restriction enzymes sites are indicated. ROD 27 and 35 are derived from integrated proviruses whereas ROD 4 is derived from a circular viral DNA. The part of the λ clones that hybridizes to the cDNA E2 is indicated below the maps: a restriction map of the ROD isolate was reconstructed from those of the three λ clones. B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; Ps, *Pst*I; Pv, *Pvu*II; S, *Sac*I; X, *Xba*I. R, L, right and left *Bam*HI arms of the λL47.1 vector. *B*, 1–11: dots corresponding to the single-stranded DNA form of M13 subclones from the HIV-1_{BRU} cloned genome. Their size and position on the HIV-1 genome, determined by sequencing[6], is shown below the figure. Dot 12 is a control containing λ phage DNA. The dot blot was hybridized in low stringency conditions (as described Fig. 1) with the complete λ ROD 4 clone as a probe, and successively washed in 2 × SSC, 0.1% SDS at 25 °C. (T_m − 42 °C), 1 × SSC, 0.1% SDS at 60 °C. (T_m − 20 °C), and 0.1 × SSC, 0.1% SDS at 60 °C. (T_m − 3 °C) and exposed overnight. A duplicate dot blot was hybridized and washed in stringent conditions (as described in Fig. 2) with the labelled λJ19 clone carrying the complete HIV-1_{BRU} genome[4]. HIV-1 and HIV-2 probes were labelled at the same specific activity (10^8 c.p.m. μg^{-1}).

Methods. DNA from the HIV-2_{ROD}-infected CEM (Fig. 2, lanes a, c) was partially digested with *Sau*3Al; the 9–15-kb fraction was selected on a 5–40% sucrose gradient, and ligated to *Bam*HI arms of the λL47.1 vector[7]. Plaques (2×10^6) obtained after *in vitro* packaging and plating on *Escherichia coli* LA 101 strain were screened *in situ* with the insert from the E2 cDNA clone. About 10 positive clones were plaque purified, and propagated on *E. coli* C600 recBC. The ROD 4, 27 and 35 clones were amplified and their DNA characterized by restriction mapping, and Southern blotting with the HIV-2 cDNA clone under stringent conditions, and *gag-pol* probes from HIV-1 used under non-stringent conditions.

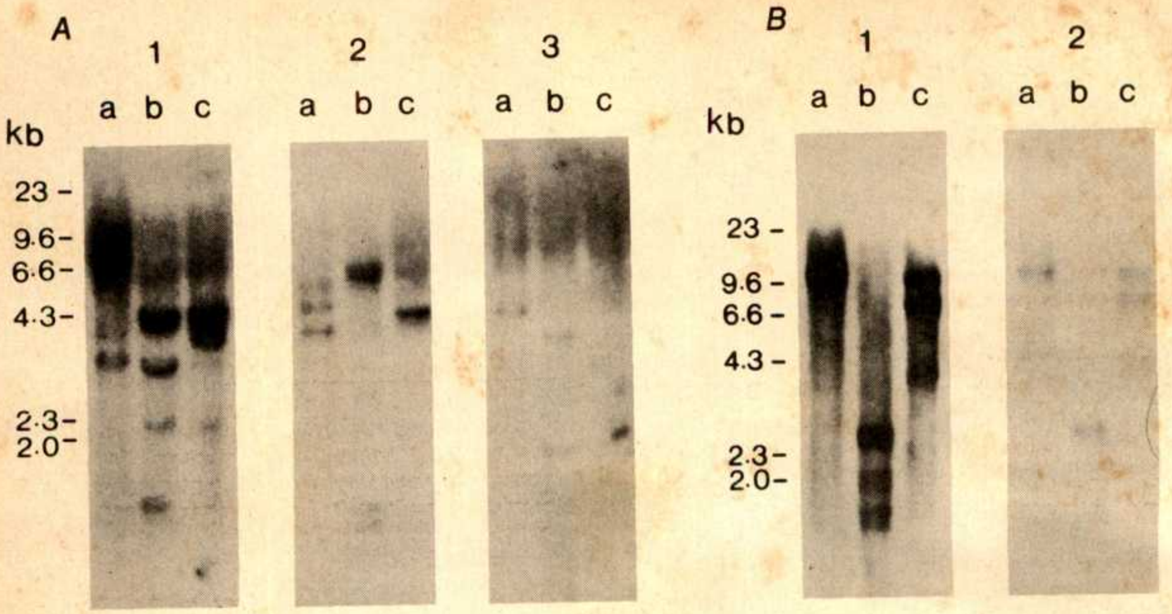


Fig. 4 Restriction map polymorphism in different HIV-2 isolates, and comparison of HIV-2 with SIV. *A*, DNA (20 μg per lane) from CEM cells infected by the isolate HIV-2_{DUL} (panel 1) or peripheral blood lymphocytes (PBL) infected by the isolates HIV-2_{GOM} (panel 2) and HIV-2_{MIR} (panel 3) was digested with: a, *Eco*RI; b, *Pst*I; and c, *Hin*dIII. Note that much less viral DNA is obtained with HIV-2 isolates propagated on PBL. Hybridization and washing were in stringent conditions, as described in Fig. 2, with 10^6 c.p.m. ml^{-1} of each of the E2 insert (cDNA) and the 5-kb *Hin*dIII fragment of ROD 4, labelled to 10^9 c.p.m. $\mu\mathrm{g}^{-1}$. *B*, DNA from HUT 78 (a T lymphoid cell line) cells infected with SIV Mm 142-83; the same amounts of DNA and enzymes were used as for *A*. Hybridization was performed with the same probe as *A*, but in non-stringent conditions, as described (Fig. 1). Washing was for 1 h in 2×SSC, 0.1% SDS at 40 °C (panel 1) and after exposure, the same filter was re-washed in 0.1×SSC, 0.1% SDS at 60 °C (panel 2). The autoradiographs were obtained after overnight exposure with intensifying screens.

infected patients, whereas the serological cross-reactivity of HIV-1 to 2 is restricted to the core proteins[2]. However SIV and HIV-2 can be distinguished by slight differences in the apparent molecular weight of their proteins[2]. In terms of nucleotide sequence, it also appears that the HIV-2 is very close to SIV. The genomic RNA of SIV (provided by Dr R. Desrosiers) can be detected in stringent conditions (Fig. 2*C*) by HIV-2 probes corresponding to the LTR and 3′ end of the genome (E2) or to the *gag* and *pol* genes (data not shown); in the same conditions, HIV-1 derived probes do not detect the SIV genome (Fig. 2*D*). In Southern blots of DNA from SIV-infected cells, a restriction pattern clearly different from HIV-2_{ROD} and other isolates is seen (Fig. 4*B*). All the bands persist after a stringent washing, although the signal is considerably weakened, indicating a sequence homology throughout the genomes of HIV-2 and SIV. It has recently been shown that baboons and macaques could be experimentally infected with HIV-2, providing an interesting animal model for the study of the HIV infection and its preventive therapy (P. Fultz and R. Desrosiers, personnal communication). Attempts to infect non-human primates with HIV-1 have only succeeded in chimpanzees, which are not a convenient model.

We have started a survey of the restriction maps of some of the HIV-2 isolates obtained by our group and from this first study it is already apparent that HIV-2, like HIV-1, undergoes restriction-site polymorphism. Figure 4*A* gives examples of such differences for three isolates, all different one from another and from the cloned HIV-2_{ROD}. It is very likely that these differences at the nucleotide level are accompanied by variations in the amino-acid sequence of the viral proteins, as in the case of HIV-1 (ref. 9).

Recently another group reported the isolation of a human retrovirus from West Africa, apparently highly related to HIV-2, and provisionally named human T-lymphotropic virus type 4 (HTLV-4, ref. 10). This virus displays tropism for $\mathrm{CD4}^+$ cells, and sera from infected patients detect all the HIV-2 proteins (F.C., unpublished data). HTLV-4 has not been isolated from AIDS patients, and therefore has been postulated to be non-pathogenic[10,11]. We feel that this apparent difference in the pathogenicity of HIV-2 and HTLV-4 is due to their means of discovery. That is, HIV-2 was isolated in a survey of AIDS cases in the West Africa, whereas HTLV-4 was found after a large scale serological screening of populations in Senegal for the presence of anti-HIV-1 antibodies and therefore primarily in healthy individual[11]. As HIV-2 and HTLV-4 infect individuals from neighbouring geographical areas, we think that they will fall into the same group of the HIV family.

Another interest in the characterization of HIV-2 is its possible contribution to the delineation of the domain of the envelope glycoprotein responsible for binding to the surface of the target cells and the subsequent internalization of the virus. This interaction was shown to be mediated by the CD4 molecule itself in the case of HIV-1 (refs 12, 13) and similar studies tend to indicate that HIV-2 uses the same receptor[2]. Thus, although there is wide divergence between the *env* genes of HIV-1 and 2, we may find small homologous domains of the envelopes of the two HIV that could represent candidate receptor-binding site. This site could constitute a target for the attempts of raising a protective immune response against this group of retroviruses.

We thank Dr F. Brun-Vézinet, Hôpital Claude Bernard, Paris, and Dr R. Desrosiers, New-England Regional Primate Research Center, for providing us with the HIV-2_{ROD} and SIV Mm 142-83 isolates, Dr Patricia Fultz, Centers for Disease Control, Atlanta, for information on experimental infections with HIV-2, Drs Pierre Sonigo and Simon Wain-Hobson for their interest in that work, and Dr Michael Emerman for critical reading of the manuscript.

Received 12 November; accepted 24 November, 1986.

1. Clavel, F. *et al. C.r. Acad. Sci.* (*Paris*) **302,** 485–488 (1986).
2. Clavel, F. *et al. Science* **233,** 343–346 (1986).
3. Weiss, R. A. in *RNA Tumor Viruses* (eds Weiss, R., Teich, N., Varmus, H. & Coffin, J.) vol. II, 405–485 (Cold Spring Harbor Laboratory, New York, 1985).

4 Alizon, M et al. Nature 312, 757-760 (1984)
5 Foley, G E et al. Cancer 18, 522-529 (1965)
6 Wain Hobson, S, Sonigo, P, Danos, O, Cole, S & Alizon, M Cell 40, 9-17 (1985)
7 Loenen, W A M & Brammar, W J Gene 10, 249-259 (1980)
8 Daniel, M D et al Science 228, 1201-1204 (1985)
9 Alizon, M, Wain-Hobson, S, Montagnier, L & Sonigo, P Cell 46, 63-74 (1986)
10 Kanki, P et al. Science 232, 238-243 (1986)
11 Barin, F et al. Lancet ii, 1387-1389 (1985)
12 Klatzmann, D et al Nature 312, 767-768 (1984)
13 Dalgleish, A et al. Nature 312, 763-767 (1984)
14 Wilbur, W J. and Lipman, D J Proc natn Acad Sci. U S A. 80, 726-730 (1983)
15 Gubler, U & Hoffman, B J Gene 25, 263-269 (1983)
16 Sanger, F, Nicklen, S & Coulsen A R Proc. natn Acad Sci. U S A. 74, 5463-5467 (1977)

A new way of enhancing the thermostability of proteases

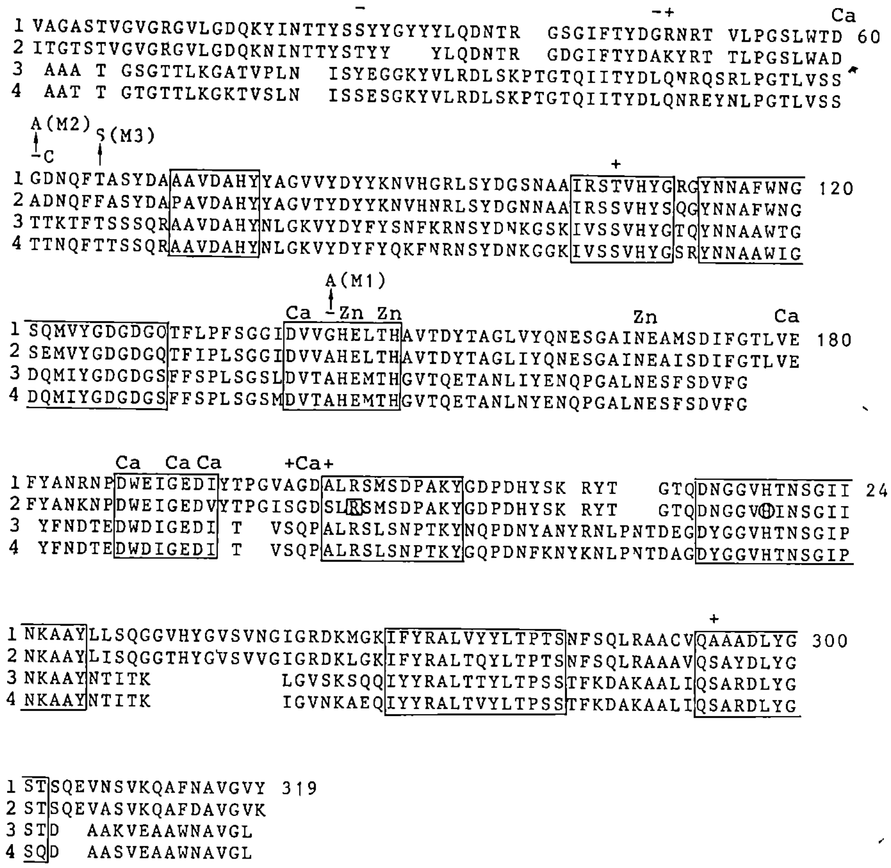
Tadayuki Imanaka, Mitsuyoshi Shibazaki & Masahiro Takagi

Department of Fermentation Technology, Faculty of Engineering, Osaka University, Yamada-oka, Suita-shi, Osaka 565, Japan

Thermostability of enzyme can be enhanced by single amino acid substitutions[1,2]. Recent advances in genetic engineering have made it possible to create novel proteins in a predictable manner where structural information for the protein is available. This 'protein engineering'[3] has already been used to enhance enzyme thermostability[4,5], but it is usually not clear which amino acid substitutions should be made We consider that the following approach should be helpful in engineering proteins with enhanced thermostability: highly conserved residues should be left unchanged; the sequences of known mesophilic and thermophilic proteins should be used to suggest the kinds of changes likely to increase thermostability[6]; and substitutions should be made that increase internal hydrophobicity and that stabilize helices for strong internal packing. We describe here the use of this approach to alter the thermostability of the thermostable neutral protease of *Bacillus stearothermophilus*, the sequence of which is known[7,8]. Surprisingly we find that a single mutation that decreases thermostability can require two mutations that increase stability to compensate for it The effects on stability are not additive, suggesting cooperativity.

Comparison of primary structures of enzymes of the same function but different origins is useful to determine the essential regions for activity, because active and/or substrate-binding sites are highly conserved in the homologous regions[9] Amino acid sequences of four neutral proteases from *B stearothermophilus*[8], *Bacillus thermoproteolyticus*[10], *Bacillus subtilis*[11] and *Bacillus amyloliquefaciens*[12] were therefore aligned and compared (Fig 1)

The amino acid sequence of thermostable neutral protease from *B stearothermophilus* was homologous (85%) with that of thermolysin from *B thermoproteolyticus* Similarly, the sequences for thermolabile neutral proteases from *B subtilis* and *B amyloliquefaciens* were also homologous (89%) In contrast, the homology between the thermostable and thermolabile enzymes was lower (~45%) (Fig 1) Nine regions were found to be conserved in all four enzymes These highly homologous regions are likely to be essential for the enzyme activity In fact, it has been reported that His 234 and Arg 206 are the active and substrate-binding sites of neutral protease (or of thermolysin), respectively[13] Zn ion is essential for the enzyme activity of thermolysin, and the three protein ligands for Zn (His 145, His 149 and Glu 169)[13] are conserved in all enzymes (Fig 1) Glu 146, which promotes the attack of a water molecule on the carbonyl carbon of the substrate[13], is also conserved in these enzymes (Fig 1)



```
                                      -                          -+             Ca
1 VAGASTVGVGRGVLGDQKYINTTYSSYYGYYYLQDNTR  GSGIFTYDGRNRT VLPGSLWTD 60
2 ITGTSTVGVGRGVLGDQKNINTTYSTYY   YLQDNTR  GDGIFTYDAKYRT TLPGSLWAD
3  AAA T GSGTTLKGATVPLN  ISYEGGKYVLRDLSKPTGTQIITYDLQNRQSRLPGTLVSS
4  AAT T GTGTTLKGKTVSLN  ISSESGKYVLRDLSKPTGTQIITYDLQNREYNLPGTLVSS

 A(M2) S(M3)
 ↑     ↑
 -C
                                                      +
1 GDNQFTASYDAAAVDAHYYAGVVYDYYKNVHGRLSYDGSNAAIRSTVHYGRGYNNAFWNG 120
2 ADNQFFASYDAPAVDAHYYAGVTYDYYKNVHNRLSYDGNNAAIRSSVHYSQGYNNAFWNG
3 TTKTFTSSSQRAAVDAHYNLGKVYDYFYSNFKRNSYDNKGSKIVSSVHYGTQYNNAAWTG
4 TTNQFTTSSQRAAVDAHYNLGKVYDYFYQKFNRNSYDNKGGKIVSSVHYGSRYNNAAWIG

                          A(M1)
                           ↑
                       Ca -Zn Zn                 Zn         Ca
1 SQMVYGDGDGOTFLPFSGGIDVVGHELTHAVTDYTAGLVYQNESGAINEAMSDIFGTLVE 180
2 SEMVYGDGDGQTFIPLSGGIDVVAHELTHAVTDYTAGLIYQNESGAINEAISDIFGTLVE
3 DQMIYGDGDGSFFSPLSGSLDVTAHEMTHGVTQETANLIYENQPGALNESFSDVFG
4 DQMIYGDGDGSFFSPLSGSMDVTAHEMTHGVTQETANLNYENQPGALNESFSDVFG

          Ca   Ca Ca      +Ca+
1 FYANRNPDWEIGEDIYTPGVAGDALRSMSDPAKYGDPDHYSK RYT    GTQDNGGVHTNSGII 240
2 FYANKNPDWEIGEDVYTPGISGDSLRSMSDPAKYGDPDHYSK RYT    GTQDNGGVHINSGII
3  YFNDTEDWDIGEDI T   VSQPALRSLSNPTKYNQPDNYANYRNLPNTDEGDYGGVHTNSGIP
4  YFNDTEDWDIGEDI T   VSQPALRSLSNPTKYGQPDNFKNYKNLPNTDAGDYGGVHTNSGIP

                                                          +
1 NKAAYLLSQGGVHYGVSVNGIGRDKMGKIFYRALVYYLTPTSNFSQLRAACVQAAADLYG 300
2 NKAAYLISQGGTHYGVSVVGIGRDKLGKIFYRALTQYLTPTSNFSQLRAAAVQSAYDLYG
3 NKAAYNTITK             LGVSKSQQIYYRALTTYLTPSSTFKDAKAALIQSARDLYG
4 NKAAYNTITK             IGVNKAEQIYYRALTVYLTPSSTFKDAKAALIQSARDLYG

1 STSQEVNSVKQAFNAVGVY 319
2 STSQEVASVKQAFDAVGVK
3 STD  AAKVEAAWNAVGL
4 SQD  AASVEAAWNAVGL
```

Fig 1 Comparison of amino acid sequences of various extracellular neutral proteases Amino acid residues are shown by single letters A, Ala, C, Cys, D, Asp, E, Glu, F, Phe, G, Gly, H, His, I, Ile, K, Lys, L, Leu, M, Met, N, Asn, P, Pro, Q, Gln, R, Arg, S, Ser, T, Thr, V, Val, W, Trp, Y, Tyr A blank indicates the absence of corresponding amino acid of this position Enzyme sources are 1, *Bacillus stearothermophilus*[8], 2, *Bacillus thermoproteolyticus*[10], 3, *Bacillus subtilis*[11], 4, *Bacillus amyloliquefaciens*[12] Homologous sequence regions are surrounded by rectangles Active and substrate-binding sites of thermolysin are indicated by ○ and □, respectively Protein ligands for Zn and Ca ions for thermolysin are indicated above the sequence Substitutions expected to enhance or reduce the thermostability of *B stearothermophilus* protease in comparison with that of thermolysin from *B thermoproteolyticus* are indicated above the sequence by + or −, respectively Vertical arrows, amino acid substitutions in *B stearothermophilus* protease

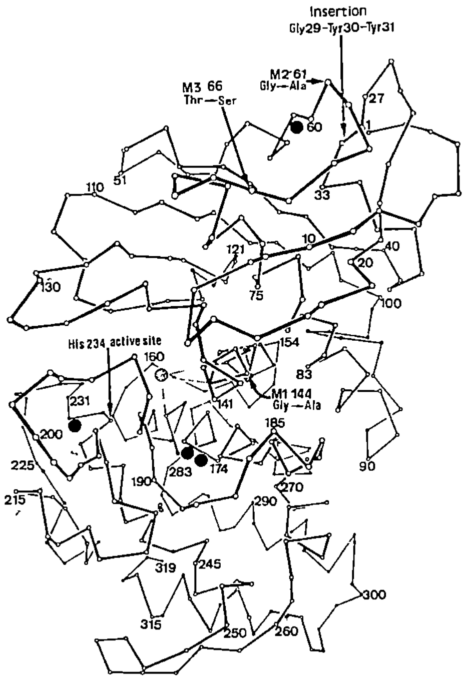


Fig 2 Three-dimensional structure of thermolysin[14] Open circles, α-carbon positions Zinc atom is drawn stippled with its three protein ligands shown diagramatically as broken lines Four calcium atoms are shown as solid circles The amino acid number from the NH_2-terminus is for *B stearothermophilus* neutral protease, originally, Gly-Tyr-Tyr at positions 29-31 are absent in thermolysin

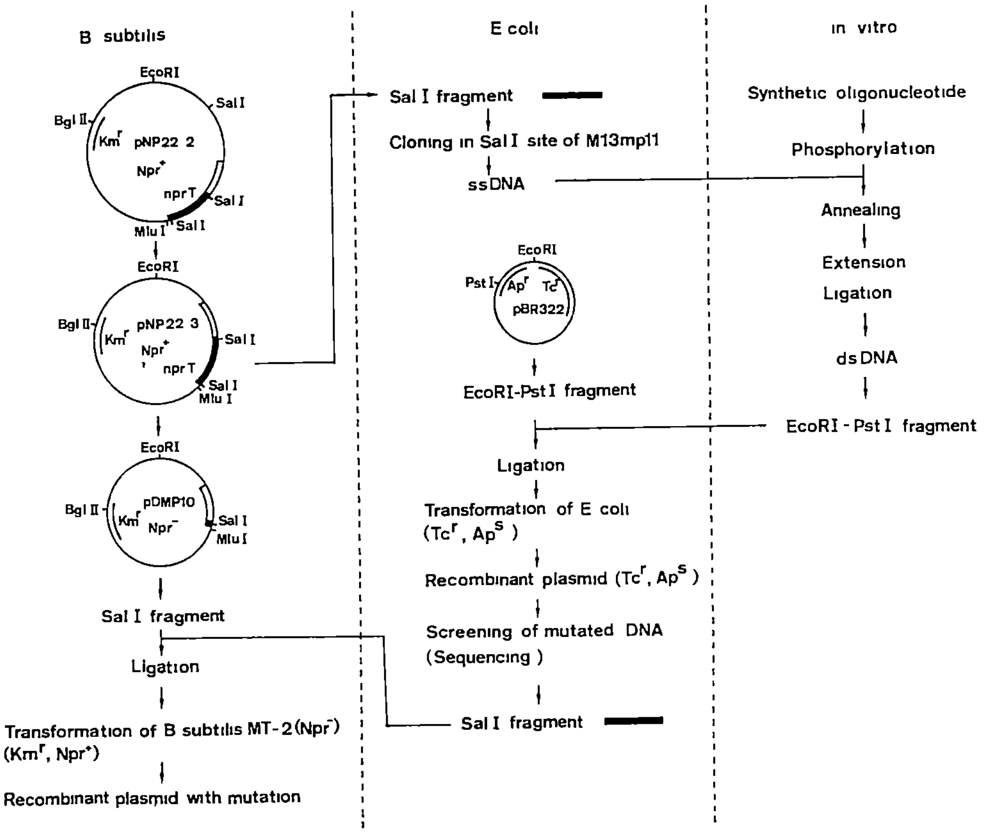


Fig 3 The scheme of site-directed mutagenesis and DNA construction The *B stearothermophilus nprT*$^+$ gene has been cloned in pNP22-2, which is a deletion derivative of pNP22-1 (ref 8) Since the plasmid contains three *Sal*I sites, one *Sal*I near an *Eco*RI site was eliminated by partial digestion of pNP22-2 with *Sal*I, digestion with Bal31, and ligation The plasmid thus obtained is designated pNP22-3 The 888 base pair (bp) *Sal*I fragment containing most of the extracellular portion of protease was isolated from pNP22-3 (Npr$^+$), and the resulting plasmid (Npr$^-$) was designated as pDMP10 The 888 bp *Sal*I fragment was cloned as described[15] in M13mp11, and single-stranded (ss) DNA was isolated The ss DNA was annealed with a synthetic oligonucleotide whose 5′ end had been phosphorylated with polynucleotide kinase After DNA extension using the primer and ligation, the double-stranded DNA was digested with both *Eco*RI and *Pst*I The *Eco*RI-*Pst*I fragment was ligated with larged *Eco*RI-*Pst*I fragment of pBR322 The ligation mixture was used to transform *Escherichia coli* Recombinant plasmids were isolated from the transformants (Tc^r Ap^s) Mutated DNA was screened by either restriction enzyme treatment (mutated primer 5′-CGTCGTG$\underline{G\overset{*}{C}GC}$ATGAGT-3′, *Hha*I for mutation M1 and 5′-GACC$\underline{GATG\overset{*}{C}}$CGACAACC-3′, *Sfa*NI for mutation M2, asterisks, indicate mismatches and underlined sequences the new restriction site) or colony hybridization[16] and the mutation was confirmed by DNA sequencing The 888 bp *Sal*I fragment with mutation was isolated from the recombinant plasmid, and ligated with *Sal*I treated pDMP10 The ligation mixture was used to transform *B subtilis* MT-2 (Npr$^-$)[7] to Km^r Npr$^+$ *B subtilis* MT-2 carrying the mutant plasmid was used to produce extracellular neutral protease

The comparison of *B stearothermophilus* protease and thermolysin revealed that the amino acid frames were completely matched except for additional three amino acids (Gly-Tyr-Tyr, 29-31) in the enzyme from *B stearothermophilus* (Fig 1) Therefore, the three-dimensional structure of *B stearothermophilus* protease would be basically similar to the known tertiary structure of thermolysin[14] shown in Fig 2 Although His 234 (active site), Arg 206 (substrate-binding site), Leu 205, Glu 193 and Asn 162 are far apart in the primary structure of the enzyme (Fig 1), these amino acid residues are near the active site in the three-dimensional structure (Fig 2) and are completely conserved in the four neutral proteases tested Therefore, alteration of amino acids in the highly conserved sequence should be avoided

We decided the position and species of amino acid to be replaced to give enhanced enzyme thermostability according to the procedure outlined The amino acid substitutions between *B stearothermophilus* protease and thermolysin were examined in the light of statistical data on amino acid substitutions which increase thermostability[6] (Fig 1) Thermolysin and *B stearothermophilus* protease contain some amino acids which promote higher thermostability in the first and the second halves, respectively Therefore, the amino acid substitution in thermostable neutral protease was confined to the first half. The substitution Gly to Ala is available for positions 47, 61 and 144 Since Gly 144 is in the α-helix which combines two domains (Fig 2), the amino acid substitution Gly 144 to Ala 144 (mutation M1) should increase the internal hydrophobicity, stabilize the α-helix and thus improve the thermostability of the protease Since it requires only the addition of a methyl group this replacement minimizes interruption of function or internal residue packing arrangements Another mutation, M3 (Thr 66 to Ser 66) should decrease thermostability of the enzyme Amino acid substitutions were performed by site-directed mutagenesis (Fig 3)

Thermostability of wild-type (WT) neutral protease of *B stearothermophilus* and thermolysin was tested at 75 °C (Fig 4*a*) Thermolysin was more thermostable than the WT protease from *B stearothermophilus* When M1 mutation (Gly 144 to Ala 144, GGG to GCG) was introduced, the M1 enzyme was found to be more thermostable than WT enzyme (Fig 4*a*, *b*) The M3 mutation (Thr 66 to Ser 66, ACC to TCC) was found to produce fairly thermolabile enzyme (Fig 4*b*) The double mutant M13 was constructed, with both the M1 and M3 mutations Surprisingly, M13 enzyme had the same thermostability as M3 When a different thermostability increasing mutation, M2 (Gly 61 to Ala 61, GGC to GCC), was added to M3, the double mutant enzyme M23 showed the same stability as M3 But when both mutations M1 and M2 were introduced into M3, the stability of the triple mutant enzyme M123 recovered to some extent compared to that of M3 (Fig 4*b*) It is interesting that two mutations are required simultaneously to enhance the thermostability of M3 enzyme

The statistical data in ref 6 were obtained from many simultaneous replacements of amino acids between thermophilic and mesophilic enzymes We attempted single amino-acid substitution to test the hypothesis and found that a single substitution does not always enhance thermostability

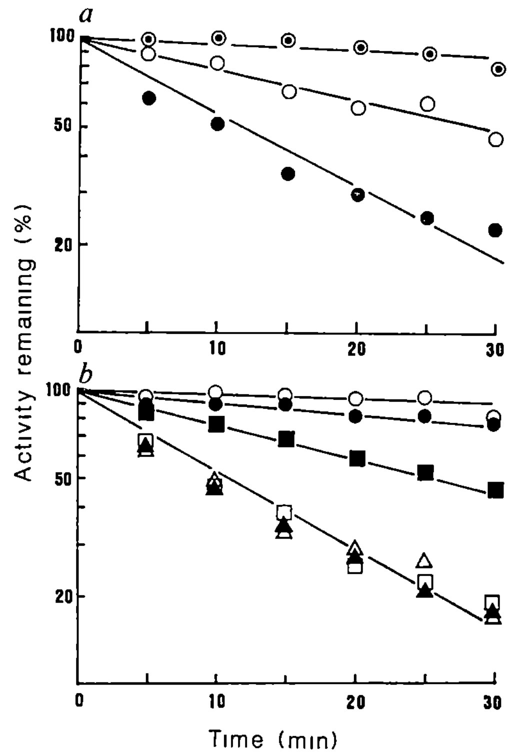


Fig 4 Thermostability of neutral proteases Extracellular protease produced by *B subtilis* carrying $nprT^+$ gene was purified by the method described previously[7] Remaining activity after heating at *a*, 75 °C, *b*, 65 °C was assayed as described earlier[7] and was expressed as the percent of original activity ○, Thermolysin, ●, wild-type enzyme (WT) from *B stearothermophilus*, ○, mutant M1, △, mutant M3, ▲, double mutant M13, □, double mutant M23, ■, triple mutant M123

The thermostability of an enzyme is influenced by many factors, including amino acid sequence, three-dimensional structure, cofactors and *p*H The *Bacillus* neutral proteases require Ca ion as the most important enzyme stabilizing factor Four Ca-binding sites were found in thermolysin (Figs 1 and 2)[14,17] Some Ca-binding sites, Asp 141, Asp 188, Glu 193 and Asp 194 are conserved for all proteases examined (Fig 1) However, a Ca-binding site, Glu 180, of thermostable proteases from *B stearothermophilus* and *B thermoproteolyticus* is deleted in the proteases from *B subtilis* and *B amyloliquefaciens* (Fig 1) This might account for the fact that the latter proteases are rather thermolabile If the amino acid sequence from Thr 177 to Phe 181 were inserted in the *B subtilis* protease, the enzyme would become more thermostable Studies on other mutations using this approach are in progress

We thank S Aiba for valuable discussion and for comments on the manuscript

Received 14 July, accepted 8 October 1986

1 Yutani, K., Ogashara, K, Sugino, Y & Matsushiro, A. *Nature* **267**, 274–275 (1977)
2 Matsumura, M, Katakura, Y, Imanaka, T & Aiba, S *J Bact* **160**, 413–420 (1984)
3 Ulmer, K M *Science* **219**, 666-671 (1983)
4 Perry, L J & Wetzel, R *Science* **226**, 555-557 (1984)
5 Wells, J A. & Powers, D B *J biol. Chem.* **261**, 6564-6570 (1986)
6 Argos, P *et al. Biochemistry* **18**, 5698-5703 (1979)
7 Fujii, M, Takagi, M, Imanaka, T & Aiba, S *J Bact* **154**, 831-837 (1983)
8 Takagi, M, Imanaka, T & Aiba, S *J Bact* **163**, 824-831 (1985)
9 Nakajima, R., Imanaka, T & Aiba, S *Appl Microbiol. Biotechnol.* **23**, 355-360 (1986)
10 Titani, K. *et al. Nature new Biol.* **238**, 35-37 (1972)
11 Yang, M Y, Ferrari, E & Henner, D J *J Bact* **160**, 15-21 (1984)
12 Vasantha, N *et al. J Bact* **159**, 811-819 (1984)
13 Kester, W R & Matthews, B W *J biol. Chem.* **252**, 7704-7710 (1977)
14 Matthews, B W *et al Nature new Biol.* **238**, 37-41 (1972)
15 Norris, K., Norris, F, Christiansen, L & Fiil, N *Nucleic Acids Res* **11**, 5103-5112 (1983)
16 Carter, P J, Winter, G, Wilkinson, A J & Fersht, A R *Cell* **38**, 835-840 (1984)
17 Matthews, B W *et al. Nature new Biol.* **238**, 41-43 (1972)

EXAFS study of the zinc-binding sites in the protein transcription factor IIIA

G. P. Diakun*, L. Fairall† & A. Klug†

* SERC Daresbury Laboratory, Daresbury,
Warrington WA4 4AD, UK
† MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

The protein transcription factor IIIA (TFIIIA) is involved in the synthesis of 5S RNA *in vitro* by RNA polymerase III. It can be isolated from *Xenopus laevis* oocytes as a 7S particle in which the protein is associated with 5S RNA. Recently it has been shown that the native particle contains 7–11 zinc atoms[1]. Analysis of the amino-acid sequence of TFIIIA revealed nine similar domains of approximately 30 amino acids, each containing two invariant pairs of histidines and cysteines, which have been implicated as possible binding sites for the zinc atoms. Other regulatory proteins with sequence homology to the zinc-binding domains of TFIIIA have now been reported[2–4]. Here, we report the results of an EXAFS (extended X-ray absorption fine structure) study of TFIIIA which shows that the coordination sphere of the zinc sites consists of two cysteine and two histidine residues.

EXAFS provides information concerning the number and type of atoms and their bond distances from a metal atom in a metalloprotein[5,6] The limitations of the technique include (1) atoms of similar atomic weight (for example oxygen and nitrogen) cannot be distinguished from one another[7], (2) outer shell distances and number become more unreliable[8], (3) the EXAFS of an atom that occupies several different sites will be an average of all the sites[9] If care is taken in the analysis the technique still offers a means of structural characterization of the local environment of the metal sites The method depends on analysing the oscillations in the absorption or fluorescence spectrum over several hundred electron volts above the metal absorption edge For biological systems where the metal content is usually low, typically ~1–6 mM, very intense X-ray sources are required and synchrotron radiation is ideal in this respect

TFIIIA was isolated as a 7S particle using the procedure of Miller *et al*[1] and the sample was kept in ice before use The protein concentration was ~2 25 mM in zinc by amino-acid analysis assuming that one zinc atom was present in each repeating unit of the protein The intensity of the fluorescence signal in the EXAFS measurement is consistent with this value, so it is probable that all nine sites are occupied by zinc atoms

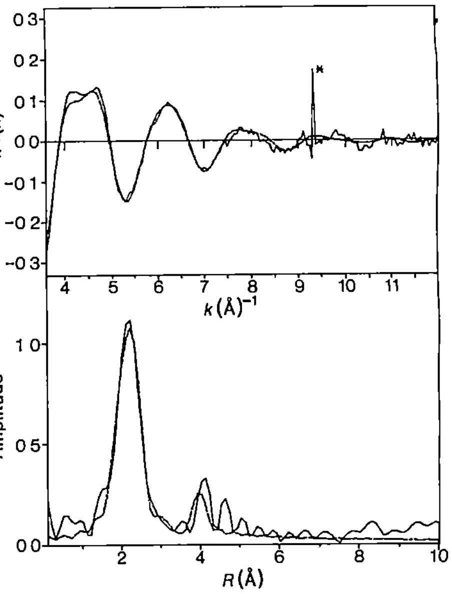


Fig. 1 Zn K-edge EXAFS and Fourier transform of transcription factor IIIA Continuous lines represent the experimental spectra, broken lines the theoretical simulation EXAFS is presented as $k\chi(k)$ plotted against k where k is in Å^{-1} It should also be noted there is a glitch in the spectrum, indicated by asterisk, this is due to other planes in the silicon crystals of the monochromator diffracting and altering the wavelength content of the beam passing through the sample[19] For the EXAFS measurements the protein solution was loaded into a teflon cell, 5 mm thick and 25×30 mm in cross-section in which a window 4×15 mm had been cut, using a syringe The sample was contained between mylar windows $\chi(k)$ is the function describing the oscillations of the EXAFS spectrum, and is multiplied by k to enhance features at high k. k is the momentum of the photoelectron given by $k = [(E - E_0)2m/\hbar^2]^{1/2}$ where m is its mass, E its energy and E_0 the energy of the absorption edge

X-ray absorption spectra were recorded over the energy range corresponding to the zinc K-edge in the fluorescence mode as the fluorescence excitation spectra with individual electronics for each of the four NaI detectors Data were collected at room temperature at the EXAFS station on the Wiggler beam line at the Daresbury Synchrotron Radiation Source operating at an energy of 2 GeV with an average current of 100 mA Harmonic contamination was minimized by using a double-crystal Si 220 monochromator[10] so that the EXAFS oscillations would be accurately measured at each position of the monochromator A total of 28 scans were collected for the TFIIIA and each spectrum was checked individually, and the output from each detector was inspected before averaging Approximately half the spectra were rejected because of odd features and jumps occurring in the data which were not reproducible and the remaining spectra were averaged to produce a single spectrum per TFIIIA with a reasonable signal to noise ratio No radiation damage was observed as no change was seen in the edge of EXAFS between spectra Any change in the local environment as a result of radiation damage would be expected to manifest itself as a change at least in the XANES (X-ray absorption near edge structure) spectrum Data analysis using the single-scattering spherical-wave method for calculating the EXAFS with phase shifts derived from *ab initio* calculations[11–13], the determination of the quality of fit and the refinement were accomplished as described previously[11]

Table 1 Parameters used to simulate the EXAFS associated with the zinc K edge of transcription factor IIIA

	Atom	No	$\sigma^2(\text{Å}^2)$*	$R(\text{Å})$†
1st nearest neighbours	N	2	0 0040	2 00
	S	2	0 0045	2 30
2nd and 3rd nearest neighbours	C	4	0 0125	3 07
	C	4	0 0035	3 88

* σ^2 is a Debye-Waller-like factor and is equivalent to $\Delta R^2_{\text{r m s}}$, the mean square deviation of an atom from its mean position
† The distance of atoms from the zinc atom

Phase shifts and the calculation procedures were checked against a model compound of known crystal structure, zinc II (1,10-phenanthroline)*bis*(4-toluenethiolate)[14] which was collected in transmission mode at the zinc K-edge (model compound supplied by C D Garner, Department of Chemistry, Manchester University) Only the primary co-ordination sphere was used to fit the model The theoretical EXAFS simulation was produced with two nitrogen atoms at ~2 10 Å and two sulphur atoms ~2 25 Å with Debye–Waller type factors (σ^2) of 0 007 and 0 006 (Å^2), compared with the crystal structure average of two nitrogen atoms at 2 112(5) Å and two sulphur atoms at 2 261(2) Å Agreement between the two methods is very good, as the error in fitting the initial shell of atoms for EXAFS is expected to be ≤ ±0 02 Å Thus the reliability of the EXAFS interpretation for the model gave us confidence in the refinement procedure and in transferring the calculated phase shifts to TFIIIA

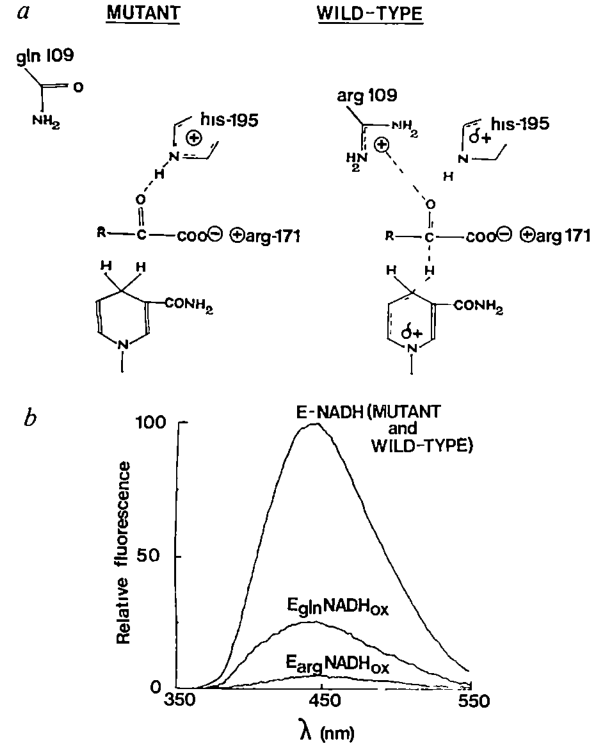


Fig 3 Changes in coenzyme fluorescence on binding oxamate to the enzyme–NADH complex *a*, Proposed structures in the active site of LDH when both NADH and oxamate are bound to the mutant and wild-type enzymes The glutamine amide is further from the substrate carbonyl than the guanidinium group The diagrams emphasize the shift in the electronic structure of the dihydronicotinamide ring in the wild-type enzyme, due to the greater polarization of the substrate carbonyl by arginine 109 Coenzyme fluorescence-emission spectra which reflect this effect[24] are shown in *b* The spectra were taken at an excitation wavelength of 340 nm and the emission was scanned between 350 nm and 550 nm The concentrations of enzyme (subunits), NADH and FBP were 15 μM, 10 μM and 5 mM, respectively and the buffer contained 20 mM triethanolamine-HCl at *p*H 6 The fluorescence enhancement (35%) and blue shift (450 nm to 440 nm at the peak) when NADH bound to the enzyme was the same for the mutant and wild-type binary complexes However, addition of a saturating concentration of oxamate (47 mM) to the mutant enzyme–NADH complex gave a quench of 75% whereas the wild-type complex (in 8 mM oxamate) showed a 95% quench (the coenzyme is five times more fluorescent in the mutant ternary complex)

intrinsic $k_H/k_D = 3$ and a rate of 750 s^{-1} would still give $k_H/k_D = 1$, with a preceding protein-isomerization rate of 250 s^{-1} Thus hydride transfer in the mutant is at least 750/0 6 = 1,200 times slower than in the wild-type enzyme Therefore compared to glutamine, arginine 109 reduces the size of the activation energy barrier for hydride transfer by more than 4 2 kcal

It has been suggested on the basis of crystal structures[1] that arginine 109 destabilizes the enzyme–NADH–pyruvate complex when the loop is closed over the active site, owing to charge repulsion between arginine 109 and the protonated histidine 195 (Fig 1) This predicts that the binding of oxamate in the ternary complex (in which the histidine is protonated and the 98–110 loop is closed[2,16,23]) should be tighter in the mutant enzyme, owing to the absence of charge repulsion from arginine-109 This suggestion is not tenable, since Table 1 shows that the binding of oxamate and pyruvate is weaker the arginine residue stabilizes the NADH-containing ternary complex by 1 6–1 9 kcal This experimental result is consistent with the observation that oxamate (isoelectronic and isosteric analogue of pyruvate) triggers loop-closure in the native enzyme[16] and thus NADH and oxamate bind more tightly when the loop is down

Table 1 Comparison of the properties of native and mutant (Arg-109 → Gln) enzymes

	Native LDH	Mutant LDH
(1) K_d NADH	1 5 μM	1 5 μM
(2) K_m pyruvate (unactivated)	2 mM	30 mM
K_m pyruvate (activated)	0 04 mM	1 0 mM
Activation factor (pyruvate)	50	30
(3) k_{cat}	250 s^{-1}	0 6 s^{-1}
(4) K_d oxamate (unactivated)	1 5 mM	25 mM
K_d oxamate (activated)	0 03 mM	1 mM
Activation factor (oxamate)	50	25
(5) pK of histidine 195 in E–NADH	6 4	6 5
(6) Addition of SO_3 to E–NAD^+		
K_d	4 μM	6 μM
k_{on}	0 94 × 10^5 M^{-1} s^{-1}	1 22 × 10^5 M^{-1} s^{-1}
(7) Thermal stability ($t_{1/2}$ 90 °C)	7 0 min	5 0 min

Unless stated, all parameters were measured at 25 °C in 20 mM triethanolamine-HCl (*p*H6) (1) Measured from NADH fluorescence anisotropy[9] (2) Determined in assay solutions containing saturating (0 2 mM) NADH Activation factor is the ratio of K_m determined in the absence and presence of 5 mM FBP (the allosteric activator)[9] (3) The maximum single-turnover rate at saturating NADH and pyruvate (experimental conditions of Fig 2) (4) Determined fluorimetrically[9] in presence and absence of 5 mM FBP (5) Determined for the bacterial enzyme-NADH complexes (6 μM sites, 5 μM NADH) in 50 mM TEA *p*H 6–8[17] (6) $K_{d(apparent)}$ at *p*H 7 2, 15 μM enzyme sites, 0 5 mM NAD^+, 5 mM FBP k_{on} Is from the variation in first-order rate constant with sulphite (75 μM to 300 μM) in stopped-flow experiments under the same conditions[18] The non-enzymatic rate[19] is 2,400 M^{-1} s^{-1} (7) Temperature stability was measured by incubating the 2 μM enzyme subunits at 90 °C in the presence of 5 mM FBP and assaying the residual enzyme activity

and the arginine residue is in the active site

If the role of arginine 109 is not to destabilize the enzyme–NADH–pyruvate complex, then an alternative means by which it stimulates the rate of hydride ion-transfer must be sought One possibility is that the residue is able to enhance the polarization of the carbonyl group of pyruvate over that achieved by histidine 195 alone (Fig 1) In their pioneering work on LDH, Winer and Schwert[24] suggested that pyruvate and oxamate are able to quench the fluorescence of enzyme-bound NADH by causing a partial transfer of the hydride ion from the nicotinamide C-4 to the substrate C-2 (Fig 3*a*) to generate in the ground state a coenzyme structure which more resembles the unfluorescent NAD^+ The extent of this partial transfer of charge, and resultant fluorescence quench, will depend on the degree of polarization of the substrate carbonyl Figure 3*b* shows the NADH fluorescence spectra in the enzyme–NADH and enzyme–NADH–oxamate complexes of the wild-type and mutant LDHs The result demonstrates that the fluorescence-quench on adding oxamate is far less pronounced in the mutant and provides experimental evidence that arginine 109 increases the polarity of the carbonyl bond of the substrate in the ground state (a function speculated upon by Grau *et al* in a proposed mechanism for pyruvate reduction[1])

An assumption of the 'mutagenesis' approach is that, in making the substitution, there are no widespread and uninterpretable effects on other regions of the protein[25] One control is to determine the three-dimensional structures of the mutant enzyme, but a rapid and more sensitive control is to measure a second catalytic function of the enzyme which is expected to be independent of the mutation LDH catalyses the addition of sulphite (SO_3^{2-}) to NAD^+ in the active site of the enzyme This reaction tests several characteristics of the active site but does not require hydride-ion transfer It depends on histidine 195 being protonated and NAD^+ being activated and orientated to facilitate covalent attachment of the sulphite ion to C-4 of the nicotinamide ring of NAD^+ through a nucleophilic attack[2,18] Table 1 shows the mutant enzyme catalyses NAD-sulphite formation slightly faster than the wild-type in contrast to the >1,200-fold reduction in the rate of hydride transfer Additional evidence that the protein structure is not significantly perturbed by the mutation is its lack of effect upon the temperature stability of the enzyme (Table 1)

In conclusion, the substitution of glutamine for arginine-109 (conserved in all 16 known LDHs and all 4 known MDHs) has changed the properties of the enzyme in a limited but mechanistically interpretable way We suggest that the binding of the anionic pyruvate at the active site neutralizes some of the positive charge in this region, triggering loop closure and exploiting the favourable hydrophobic contacts[2 16] made between the loop and the body of the protein This brings arginine 109 close to the substrate carbonyl (an example of a substrate-induced conformational change) Our data show that the relative enzyme-substrate ground state stabilization of the arginine 109 containing complex, over that containing glutamine, is 1 6–1 9 kcal, while the relative transition state stabilization is 4 2 kcal In the ground state the polarization of the carbonyl bond of the substrate appears to be enhanced by arginine 109 (Fig 3*b*) The greater effect of the loss of arginine 109 on the transition state stability is consistent with the carbonyl bond becoming more polarized as a negative charge develops on the carbonyl oxygen when C-2 goes from sp^2 to sp^3 hybridization

We thank Professor M Buehner for his discussions of the crystallographic structure of *L casei* LDH and the SERC for a project grant (protein engineering), personal fellowship (A R C) and studentship (D B W)

Received 6 May, accepted 3 November 1986

1 Grau U M , Trommer, W E & Rossmann, M G *J molec Biol* **151**, 289-307 (1981)
2 Parker, D M & Holbrook, J J *Pyridine Nucleotide Dependent Dehydrogenases* 485-502 (ed Sund, H) (de Gruyter, Berlin, 1977)
3 Birktoft, J J , Fernley, R T , Bradshaw, R A & Banaszak, L J *Proc natn Acad Sci U S A* **79**, 6166-6170 (1982)
4 Adams, M J *et al J molec Biol* **41**, 159-188 (1969)
5 Adams, M J *et al Proc natn Acad Sci U S A* **70**, 1968-1972 (1973)
6 White, J L *et al Molec Biol* **102**, 759-779 (1976)
7 Barstow, D A *et al Gene* **46**, 47-55 (1986)
8 Wirz B , Suter, F & Zuber, H *Hoppe-Seyler s Z physiol Chem* **364**, 893-909 (1983)
9 Clarke, A R , Atkinson, T , Campbell, J W & Holbrook, J J *Biochim biophys Acta* **829**, 387-396 (1985)
10 Winter, G , Fersht, A R , Wilkinson, A J , Zoller, M & Smith, M *Nature* **299**, 756-758 (1982)
11 Amann E Brosius, J & Ptashne M *Gene* **25**, 167-178 (1983)
12 Gibson, T J thesis, Univ Cambridge (1984)
13 Clarke, A R , Waldman, A D B , Munro, I & Holbrook, J J *Biochim biophys Acta* **828**, 375-379 (1985)
14 Atkinson, T *et al* in *Bioactive Microbial Products* (ed Stowell, J D) Vol 3, 27-43 (Academic, London, 1986)
15 Holbrook, J J , Liljas, A , Steindel, S J & Rossmann, M G in *The Enzymes* (ed Boyer, P D), 3rd Edn Vol 11, 191-292 (Academic, New York, 1975)
16 Parker, D M , Jeckel, D & Holbrook, J J *Biochem J* **201**, 465-471 (1982)
17 Holbrook, J J & Stinson, R A *Biochem J* **131**, 739-748 (1973)
18 Parker, D M Lodola, A & Holbrook, J J *Biochem J* **173**, 959-967 (1978)
19 Johnson, S L & Smith, K W *Biochemistry* **15**, 553-559 (1976)
20 Holbrook, J J & Gutfreund H *FEBS Lett* **31**, 157-169 (1973)
21 Klinman, J P *Adv Enzym relat Areas molec Biol* **46**, 415-493 (1978)
22 Lodola, A , Shore, J D , Parker, D M & Holbrook, J J *Biochem J* **175**, 987-998 (1978)
23 Clarke, A R , Waldman, A D B , Hart, K W & Holbrook, J J *Biochim biophys Acta* **829**, 397-407 (1985)
24 Winter, A D & Schwert, G W *J biol Chem* **234**, 1155-1161 (1959)
25 Ackers, G K & Smith F R *A Rev Biochem* **54**, 597-629 (1985)

Corrigenda

A complimentary DNA clone for a macrophage-lymphocyte Fc receptor

Victoria A. Lewis, Terry Koch, Helen Plutner & Ira Mellman
Nature **324**, 372–375 (1986)

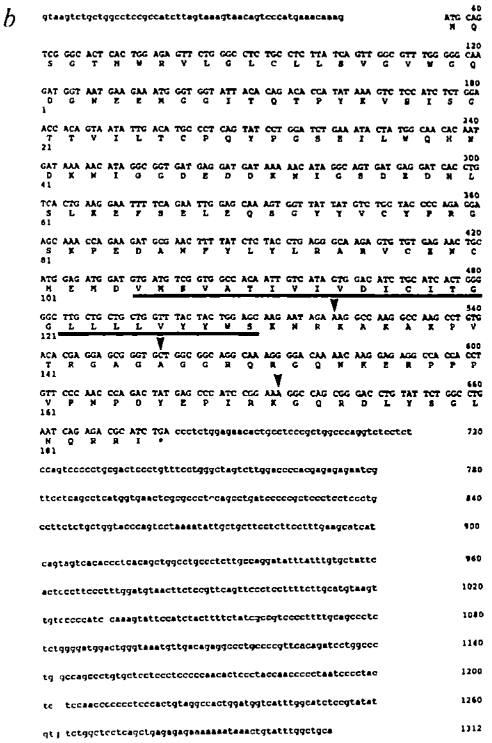
A LATE correction to Fig 1*b* in this letter was omitted The nucleotide numbering for pFcRAb should read 805–1421, not 805–1385 as shown The figure appears correctly on reprints The insert is of 6/6 base pairs, not 580 as in paragraph 2 of the letter

---

Isolation of cDNA clones encoding the 20K non-glycosylated polypeptide chain of the human T-cell receptor/T3 complex

Daniel P. Gold, Jennifer M. Puck, Carolyn L. Pettey, Mildred Cho, John Coligan, James N. Woody & Cox Terhorst
Nature **321**, 431–434 (1986)

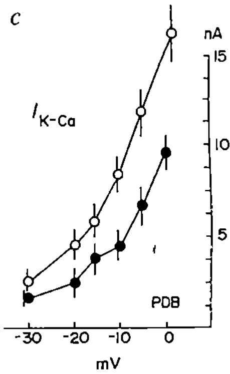
THREE sequencing errors have been found in the human T3-ε nucleotide sequence published in Fig 3*b* These errors have been corrected in the sequence shown below at the positions indicated by vertical arrows The putative transmembrane region is underlined The predicted amino acid sequence ends after residue 185 instead of the reported residue 211

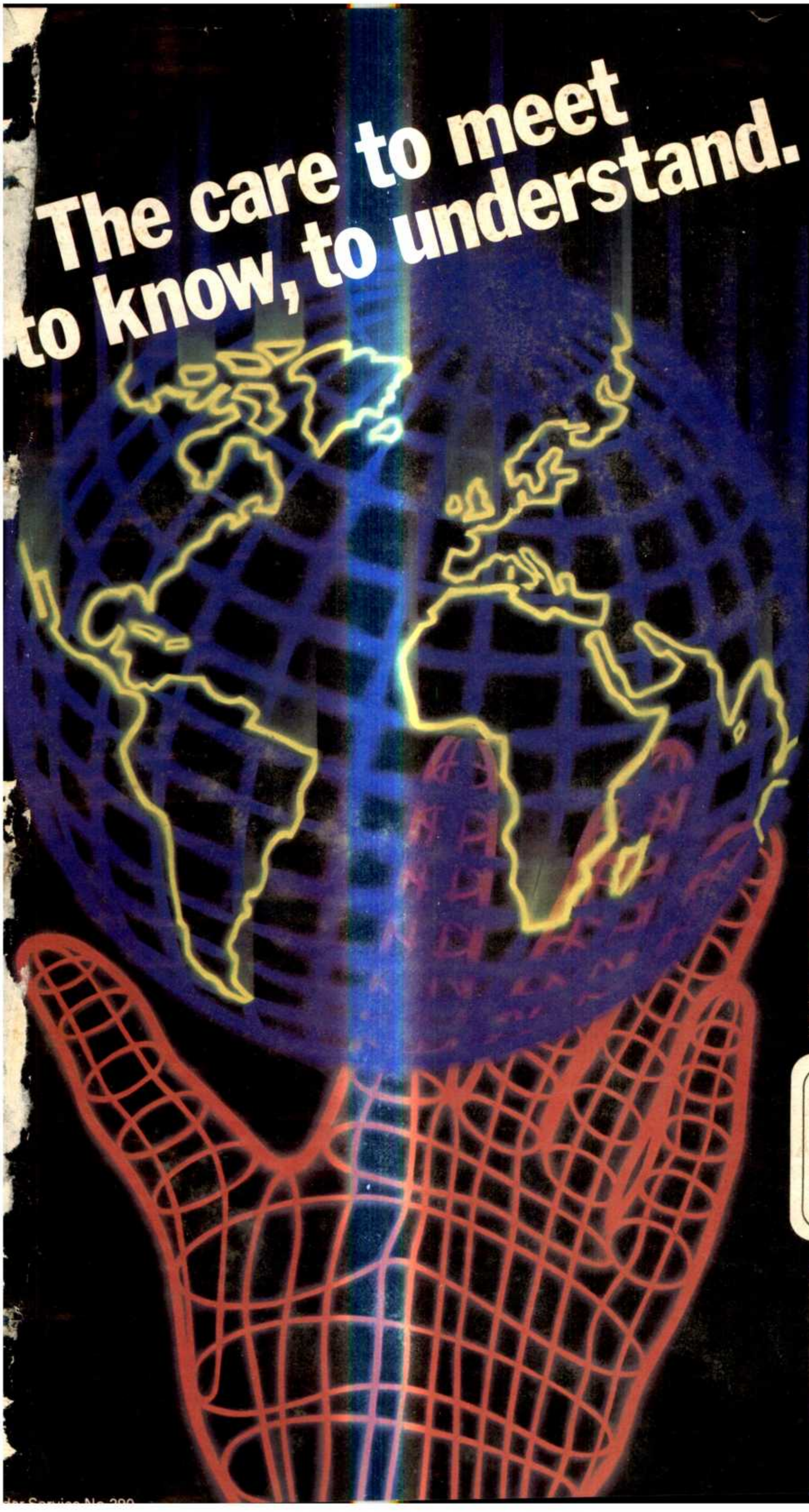


Protein kinase C activation induces conductance changes in *Hermissenda* photoreceptors like those seen in associative learning

J. Farley & S. Auerbach
Nature **319**, 220–223 (1986)

PANEL *c* in Fig 3 of this letter was incorrectly drawn With the ordinate scale corrected the figure reads





Ares-Serono Symposia is an independent foundation, created 1971, to promote scientific research in all disciplines which contribute towards improving human health.
This aim is pursued by means of congresses, courses, seminars and specialized studies.

Some of the international meetings scheduled for 1987 are:

Herpes and Papilloma Viruses
Milan, Italy / March 26-27
Scientific Organization: F. De Palo (I), F. Rilke (I) and H. Zur Hausen (D)
IV Pan American Congress of Androlog
São Paulo, Brazil / May 4-6
Scientific Organization: A. Negro-Vilar (USA) and M. P. De Castro (BZ)
Inhibin - Non-Steroidal Regulation of Follicle Stimulating Hormone Secretio
Tokyo, Japan / May 21-22
Scientific Organization: H. Burger (Aus) an M. Igarashi (J)
Development and Function of the Reproductive Organs
Turku, Finland / June 10-12
Scientific Organization: M. Parvinen (SF)
IV Colloquium of the European Pineal Study Group
Modena, Italy / August 31 - September 4
Scientific Organization: G.P. Trentini (I), A. Oksche (D) and P. Pévet (F)
Cell-to-Cell Communication in Endocrinology
Florence, Italy / October 8-9
Scientific Organization: L. Martini (I), M. Serio (I) and C.W. Bardin (USA)
Differentiation Therapy for Cancer
Tucker's Town, Bermuda / October 23-2
Scientific Organization: G.B. Rossi (I), F. Takaku (J) and S. Waxman (USA)

ARES
SERONO
SYMPOSIA

Do you know all the advantages that the Ares-Serono Symposia Congress Card offers? Please fill in this form and mail to Ares-Serono Symposia for information.

swissair Official Carrier

I would like to receive information about:
☐ Ares-Serono Symposia Congress Card
☐ All the above Congresses
☐ In particular the Congress on ....

....

....

☐ Ares-Serono Symposia Publications

Name ....

Address ....

Institution ....

Please send to Ares-Serono Symposia

der Service No. 280

cDNA Synthesis Kit from Boehringer Mannheim

the time & money saving alternative

cDNA Synthesis Kit, Cat. No. 1013 882
contains all reagents ready for use
for the transcription of 50 µg poly (A)$^+$RNA into cDNA

Get the detailed protocol from your local representative or

boehringer mannheim

Boehringer Mannheim GmbH
Biochemica
P.O. Box 310120
D-6800 Mannheim
W. Germany

Reader Service No.1